

IADIS MULTI CONFERENCE  
ON COMPUTER SCIENCE  
AND INFORMATION SYSTEMS

Rome, ITALY  
20 -26 JULY



Proceedings of the  
IADIS International Conference  
Interfaces and  
Human Computer Interaction 2011

EDITED BY  
Katherine Blashki

**IADIS INTERNATIONAL CONFERENCE**

**INTERFACES AND HUMAN  
COMPUTER INTERACTION  
2011**

**part of the**

**IADIS MULTI CONFERENCE ON COMPUTER SCIENCE AND  
INFORMATION SYSTEMS 2011**



**PROCEEDINGS OF THE  
IADIS INTERNATIONAL CONFERENCE  
INTERFACES AND HUMAN  
COMPUTER INTERACTION 2011**

**part of the**

**IADIS MULTI CONFERENCE ON COMPUTER SCIENCE AND  
INFORMATION SYSTEMS 2011**

**Rome, Italy**

**JULY 24 - 26, 2011**

**Organised by  
IADIS**

**International Association for Development of the Information Society**

Copyright 2011

IADIS Press

All rights reserved

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Permission for use must always be obtained from IADIS Press. Please contact [secretariat@iadis.org](mailto:secretariat@iadis.org)

Interfaces and Human Computer Interaction Volume Editor:  
Katherine Blashki

Computer Science and Information Systems Series Editors:  
Piet Kommers and Pedro Isaías

Associate Editors: Luís Rodrigues

ISBN: 978-972-8939-52-6

# TABLE OF CONTENTS

FOREWORD	xiii
PROGRAM COMMITTEE	xv
KEYNOTE LECTURES	xix
KEYNOTE PAPERS	xxi

## FULL PAPERS

A COMPARISON OF IN-LAB AND SYNCHRONOUS REMOTE USABILITY TESTING METHODS: EFFECTIVENESS PERSPECTIVE <i>Obead Alhadreti, Pam Mayhew and Majed Alshamari</i>	3
A COMPARATIVE ANALYSIS OF THE MEANING OF 'LEARNABILITY' FOR CHILD AND ADULT USERS <i>Bester Chimbo, Helene Gelderblom and Ruth de Villiers</i>	10
A SEMIOTIC PERSPECTIVE TO WEB USABILITY: AN EMPIRICAL CASE STUDY <i>Muhammad Nazrul Islam</i>	19
DEVELOPING USER INTERFACES FOR MONITORING SYSTEMS IN BUILDINGS <i>Szu-Cheng Chien, Robert Zach and Ardeshir Mahdavi</i>	29
READING AS PLAYING: A NEW TUTORING MULTIMEDIA TOOL FOR CHILDREN WITH TEXT COMPREHENSION PROBLEMS <i>Michela Carlini, Tania Di Mascio and Rosella Gennari</i>	37
USING GOOGLE ANALYTICS TO EVALUATE IMPROVEMENTS IN THE USABILITY AND PERFORMANCE OF E-COMMERCE WEBSITES <i>Layla Hasan</i>	45
TAILORABLE FLEXIBILITY: MAKING END-USERS AUTONOMOUS IN THE DESIGN OF ACTIVE INTERFACES <i>Federico Cabitza, Iade Gesso and Stefano Corna</i>	53
PERSONIFICATION OF TOPICS WITH CONVERSATIONAL AGENTS <i>Jean-Paul Sansonnet, François Bouchet and William A. Turner</i>	61
EVALUATION OF GESTURAL INTERACTION WITH AND WITHOUT VOICE COMMANDS <i>Joana Neca and Carlos Duarte</i>	69

MULTIMODAL INTERACTION FOR USERS WITH AUTISM IN A 3D EDUCATIONAL ENVIRONMENT <i>Alessandro Trivilini, Licia Sbattella and Roberto Tedesco</i>	77
GENERAL AND AUTOMATIC EFFICIENCY TESTING FRAMEWORK FOR JAVA <i>Narin Işık and Kerem Önal</i>	86
INTERACTION ERROR BASED VIEWPOINT ESTIMATION FOR CONTINUOUS PARALLAX ERROR CORRECTION ON INTERACTIVE SCREENS <i>Bastian Migge, Andreas Kunz and Tim Schmidt</i>	93
EXTENDING A USER OBSERVATION PROTOCOL TO ACCOUNT FOR PSYCHOLOGICAL TRAITS <i>Yuska P. C. Aguiar, Maria de F.Q. Vieira, Edith Galy, Jean-Marc Mercantini and Charles Santoni</i>	105
LAYOUT TEMPLATES - LET USERS RULE USER INTERFACES <i>Florian Haag, Christiane Taras and Thomas Ertl</i>	113
USABILITY EVALUATION METHOD <i>Junko Shirogane, Yuichiro Yashita, Hajime Iwata and Yoshiaki Fukazawa</i>	121
TABLETOP USER INTERFACE FOR NAVIGATION IN VIRTUAL ENVIRONMENTS <i>Tatu Harviainen, Hanna Koskinen and Jari Laarni</i>	129
EXPOSING REPETITIVE SCANNING IN EYE MOVEMENT SEQUENCES WITH T-PATTERN DETECTION <i>Marcus Mast and Michael Burmester</i>	137
A USABILITY STUDY OF MICROSOFT OFFICE 2007 AND MICROSOFT OFFICE 2003 <i>Edgard Velez and Nestor J. Rodriguez</i>	146
EVALUATION OF DIFFERENT INTERACTION TECHNIQUES FOR TOUCH DEVICES <i>Michael Raschke, Tanja Blascheck, David Bold, Dominik Muhler, Thomas Ertl and Thomas Schlegel</i>	153
INTERACTION DECONSTRUCTION METHOD FOR USABILITY REQUIREMENTS ENGINEERING <i>Lamminen Juha, Rousi Rebekah and Saariluoma Pertti</i>	161
COGNITIVE CHUNKS EXTRACTED FROM EYE-TRACKING RECORDS OF WEB PAGE READERS, USING ISOLATED SACCADES AS A DELIMITER <i>Noriyuki Matsuda and Haruhiko Takeuchi</i>	169
INTERACTIVE ART SYSTEM FOR MULTIPLE USERS BASED ON TRACKING HAND MOVEMENTS <i>Byunghyung Kim, Danna Gurari, Hough O'Donnell, Margrit Betke</i>	177
FEATURE EXTRACTION BASED ON EMPIRICAL MODE DECOMPOSITION AND BAND POWER APPROACHES FOR MOTOR IMAGERY TASKS CLASSIFICATION <i>Dalila Trad, Tarik Al-ani, Eric Monacelli, Stéphane Delaplace and Mohamed Jemni</i>	185
TOWARDS AN ONTOLOGY-BASED COMMUNITY OF AGENTS FOR PERSONALISATION OF SERVICES FOR DISABLED STUDENTS <i>Julius T. Nganji and Mike Brayshaw</i>	193

SUPPORTING DOWN PEOPLE IN CUTTING OPERATIONS THROUGH HAPTIC TECHNOLOGY	201
<i>Mario Covarrubias, Monica Bordegoni and Umberto Cugini</i>	
INTERPRETATION OF AMBIGUOUS IMAGES INSPECTED BY THE STICKGRIP DEVICE	209
<i>Tatiana V. Evreinova, Grigori Evreinov and Roope Raisamo</i>	
A SYSTEMATIC EVALUATION OF MOBILE SPREADSHEET APPS	217
<i>Derek Flood, Rachel Harrison, Clare Martin and Kevin McDaid</i>	
TOWARDS EVALUATING HUMAN-INSTRUCTABLE SOFTWARE AGENTS	225
<i>Robert D. Grant, David DeAngelis, Dan Luu, and Dewayne E. Perry and Kathy Ryall</i>	
QUALITATIVE RESULTS OF AN EXPERIMENT WITH ADJUSTABLE INTERACTIVE RINGS FOR IDTV	233
<i>Leonardo Cunha de Miranda, Heiko Horst Hornung and M. Cecilia C. Baranauskas</i>	
DESIGN CREATION SYSTEM BASED ON DESIGNER'S KANSEI AND CONSUMER'S PREFERENCE	241
<i>Hideki Aoyama and Kentaro Ohshima</i>	
COUPLED MOBILE DISPLAYS FOR CREATIVITY TECHNIQUE BASED PROBLEM SOLVING PROCESSES IN TABLETOP ENVIRONMENTS	249
<i>Marc René Frieß and Martin Kleinhans</i>	
ATTRACTIVENESS OF SOCIAL PLATFORMS: AN INVESTIGATION OF STUDIVZ	257
<i>Waltraud Ilmberger, Martin Schrepp and Theo Held</i>	
TOOL SUPPORT IN INTEGRATING TASK MODELING AND ACCESS CONTROL SPECIFICATION	265
<i>Birgit Bomsdorf and Annika Paulheim</i>	
MODEL-DRIVEN ENGINEERING OF BEHAVIORS FOR USER INTERFACES IN MULTIPLE CONTEXTS OF USE	273
<i>Efrem Mbaki Luzayisu and Jean Vanderdonck</i>	
EMOTIONAL WEB-BASED DESIGN: THE CONCEPTS OF EMOTIONAL EXPERIENCE AND EMOTIONAL EXPRESSION	283
<i>Zacharias Lekkas, Nikos Tsianos, Panagiotis Germanakos, Constantinos Mourlas and George Samaras</i>	
A STUDY OF A NAVIGATION SYSTEM THAT INDUCES TOURISTS TO VISIT SIGHTSEEING SPOTS AGAIN VIA A FEELING OF REGRET	291
<i>Masaki Masuda and Yoshio Nakatani</i>	
PROCESS-INDUCED DECISION COSTS ON SEQUENTIAL VALUE JUDGMENTS	299
<i>Mei-Chun Wu and Feng-Yang Kuo</i>	



## SHORT PAPERS

TESTING NEW SCENARIOS FOR DISTANCE LEARNING, BASED ON THE INTRODUCTION OF NEW DEVICES <i>Victor Ximenes and Gabriel Pontual</i>	311
USING VITAL-SENSOR IN TRACKING USER EMOTION AS A CONTEXTUAL INPUT FOR MUSIC RECOMMENDATION SYSTEM <i>Nguyen Thuy Le, Jin Nakazawa, Kazunori Takashio and Hideyuki Tokuda</i>	316
CULTURA: A TWO CULTURAL MODEL FOR UNDERSTANDING HUMAN BEHAVIOR ON THE INTERNET <i>Santosh Kumar Kalwar</i>	321
3D GESTURE RECOGNITION SYSTEM BASED ON DEPTH CAMERA <i>Yangkeun Ahn and Jiman Hong</i>	327
ENHANCING ZONE OF PRODUCT DESIGN WITH WIKI PRACTICE <i>Nikolay Yaremko</i>	332
TOUCH-BASED INTERACTION: A PROPOSAL FOR AAL <i>Bravo, J., Hervás, R. and Fontecha, J.</i>	337
INTERACTIVE SONIFICATION OF MEDICAL DATA <i>Stefan Kovar and Margit Pohl</i>	341
VISUALIZING THE IMPACT OF CONTENT-BASED SIMILARITY AND SPATIAL DISTANCE ON BOOK RECOMMENDATIONS <i>Sönke Knoch and Alexander Kröner</i>	345
SELECT’N’LABEL – DYNAMIC BOUNDARY LABEL PLACEMENT FOR MOBILE TOUCH DEVICES <i>Dominik Magin and Daniel Steffen</i>	351
TANGIBILITY AND HUMAN-COMPUTER INTERACTION: AN ALTERNATIVE APPROACH TO AFFORDANCE <i>Jean Caelen, Mélanie Becker and Anne Pellegrin</i>	355
SHOW-AND-TELL PLAY-IN: COMBINING NATURAL LANGUAGE WITH USER INTERACTION FOR SPECIFYING BEHAVIOR <i>Michal Gordon and David Harel</i>	360
EVALUATING EFFECTS OF VISUAL APPEARANCE OF VIRTUAL HUMANS IN E-COMMERCE <i>Rainer Blum and Malcolm Clarke</i>	365
A FACIAL ANIMATION FRAMEWORK WITH EMOTIVE/EXPRESSIVE CAPABILITIES <i>G.Riccardo Leone and Piero Cosi</i>	369
ARABIC USABILITY GUIDELINES <i>Perihan Elbaz</i>	374
COMMUNICATION SUPPORT BETWEEN PATIENTS DURING TREATMENT VIA USE OF MEMORIES <i>Takahiro Kihara and Yoshio Nakatani</i>	379

SMART BEER COASTER	384
<i>Leonardo Rocha, Augusto Lohmann, André Braz, Breno Bitarello and Fernando Reiszal</i>	
CONCURRENT STUDY ON PERCEPTION AND BEHAVIOR OF DRIVER IN SIMULATED DRIVING	389
<i>Clare Martin, Derek Flood and Rachel Harrison</i>	
OPTIMAL POSITIONS OF ADVERTISEMENTS ON NEWS WEBSITES FOCUSING ON THREE CONFLICTING OBJECTIVES	394
<i>Noriyuki Muramatsu, Keiki Takadama and Maki Sakamoto</i>	
IMPROVEMENT OF DETECTION PROBABILITY OF INSPECTION PANEL BY MULTIPLE MODARITIES	399
<i>Hidetoshi Nakayasu, Masao Nakagawa, Tetsuya Miyoshi and Patrick Patterson</i>	
SUPPORTING THE PARTICIPATION IN SOCIAL LIFE OF FRAIL OLDER ADULTS IN A ONE-DAY CLINIC	404
<i>Suzanne Kieffer, Christian Swine and Benoît Macq</i>	
A MULTIDISCIPLINARY REVIEW: IMPACTS ON DESIGN FOR AGE-RELATED COGNITIVE IMPAIRMENTS	409
<i>O. K. Burmeister</i>	
A CONTROL AND TRAINING SYSTEM FOR REHABILITATION DEVICES	413
<i>Patrick Patterson and Eric Sinzinger</i>	
TR-AID ADVANCED SEARCH: TAXONOMY DESIGN FOR DATA VISUALIZATION AND INFORMATION RETRIEVAL INTERFACE	417
<i>Massimiliano Gusmini and Anjula Garg</i>	
A COMBINED AGILE METHODOLOGY FOR THE EVALUATION OF WEB ACCESSIBILITY	423
<i>Firas Masri and Sergio Lujan Mora</i>	
DEVELOPING E-LEARNING SYSTEMS FOR DEAF EDUCATION: APPLICATION OF EMBODIEMENT THEORIES	429
<i>Maria Tagarelli De Monte, Daniele Capuano, Maria Roccaforte and Elena Tomasuolo</i>	
A PROTOCOL FOR EVALUATING MOBILE APPLICATIONS	434
<i>Clare Martin, Derek Flood and Rachel Harrison</i>	
THE EFFECT OF THE DIRECTION OF A GRADIENT ON HAPTIC POINTING IN A 2D VIRTUAL ENVIRONMENT	438
<i>Ludovic Potier, Olivier Gapenne, Xuan Liu and Dominique Aubert</i>	
TAMULATOR: A TOOL TO MANAGE TASK MODEL-BASED USABILITY EVALUATION IN DEVELOPMENT ENVIRONMENTS	442
<i>Shah Rukh Humayoun, Yael Dubinsky, Eli Nazarov, Assaf Israel and Tiziana Catarci</i>	
ON THE FACTORS OF THE WEBPAGE NAVIGATIONAL PERFORMANCE	447
<i>Wen-Te Chang, Kuo-Chen Huang and Chan-Li Lin</i>	

## REFLECTION PAPERS

SENSOR-BASED ESTIMATION OF PSYCHOLOGICAL STATES <i>Javier Eiguez Guevara and Hiroyuki Umemuro</i>	455
PLAYING LIKE A GIRL: GENDERED GAMING AND GIRL GAMERS <i>Pirita Ihamäki</i>	459
ENGINEERING ANONYMITY TO REDUCE AGGRESSION ONLINE <i>Michail Tsikerdekis</i>	463

## POSTERS

STUDY ON DICHROMATISM'S COLOR APPEARANCE TO LED LIGHTS <i>Mamoru Takamatsu, Yoshio Nakashima, Hiroki Fujita and Kazuhiro Sawa</i>	471
COLOR SENSITIVITY TO LED LIGHTS IN THE ELDERLY <i>Yoshio Nakashima, Mamoru Takamatsu, Hiroki Fujita, Kazuo Sendai and Santarou Nakajima</i>	474
AN INTRODUCTORY STUDY ON PLACEMENT METHOD OF BACKGROUND MUSIC FOR SHOPPING WEBSITE <i>Chien-Jung Lai, Wen-Ming Wu, Ya-Ling Wu and You-Shyang Chen</i>	477
PETRI NET BASED HCI SPECIFICATION AND CONTROL FOR VIRTUAL MAINTENACE TRAINING <i>Wen-Yang Feng, Chin-Jung Chao, Feng-Yi Tseng and Tien-Lung Sun</i>	480
VISUAL AND MANUAL PERFORMANCE OF HUMAN COMPUTER INTERACTION IN VIRTUAL ENVIRONMENT <i>Chiuhsiang Joe Lin, Hung-Jen Chen and Tien-Lung Sun</i>	483
HUMAN PULSE WAVE MEASUREMENT BY MEMS-ECM SENSOR <i>Shusaku Nomura, Yasushi Hanasaka, Masako Hasegawa-Ohira, Tadashi Ishiguro and Hiroshi Ogawa</i>	486
EVALUATION OF WCAG 2.0 SUCCESS CRITERIA ON LUMINANCE CONTRAST RATIO BASED ON AGE-RELATED SPECTRAL EFFICACY <i>Tetstuya Fujita, Yoshio Nakashima and Mamoru Takamatsu</i>	489
DEVELOPMENT OF RESIDENTIAL MONITORING SYSTEM BY MEASURING FIXTURES ACCELERATION <i>Macky Kato, Yoshie Shimodaira and Takeshi Sato</i>	492
LEVERAGING OCR TECHNIQUE IN VIRTUAL KEYBOARD IMPLEMENTATION <i>Jure Mornar and Andrina Granić</i>	495
DESIGN OF A FOLDING WORKBENCH AND A VIRTUAL REALITY SIMULATION SYSTEM USING IT <i>Hyun Suk Lee, Bon-Wook Gu, Jongwon Back, and Tae-Jeong Jang</i>	499
A STUDY ON CATEGORY TYPE OF ONLINE SHOPPING MALL <i>Kristen Koeun Nam, Kyle Kim and Sungmin Hong</i>	503

NONVERBAL COMMUNICATION TOOL VIA USE OF PERCUSSION  
PERFORMANCES AND ITS EFFECTIVENESS 506

*Yuji Kawase and Yoshio Nakatani*

VR INTERACTION TOOLS FOR MOTOR IMPAIRMENT SIMULATION 509

*Marco Fontana and Marcello Carrozzino*

## **DOCTORAL CONSORTIUM**

SENSORY FEEDBACK FOR MOTOR PERFORMANCE IN VIRTUAL  
ENVIRONMENTS ACROSS THE LIFESPAN 515

*Patrick J. Grabowski and Andrea H. Mason*

NEURAL BASIS OF COMMUNICATION OF AUDITORY BCIs 520

*Angela Riccio, Luca Simione, Donatella Mattia, Febo Cincotti and Marta Olivetti Belardinelli*

AUTHOR INDEX



# FOREWORD

These proceedings contain the papers of the IADIS International Conference on Interfaces and Human Computer Interaction 2011, which was organised by the International Association for Development of the Information Society in Rome, Italy, 24 – 26 July, 2011. This conference is part of the Multi Conference on Computer Science and Information Systems 2011, 20 - 26 July 2011, which had a total of 1402 submissions.

The IADIS Interfaces and Human Computer Interaction (IHCI) 2011 conference aims to address the main issues of concern within Interface Culture and Design with a particular emphasis on the affective aspects of design, development and implementation of interfaces and the generational implications for design of human and technology interaction. This conference aims to explore and discuss innovative studies of technology and its application in interfaces and welcomes research in progress, case studies, practical demonstrations and workshops in addition to the traditional submission categories.

Submissions were accepted under the following topics:

- Affective User-Centred Analysis, Design and Evaluation
- The Value of Affective Interfaces / Systems / Application / Interaction
- Generational differences and technology design
- Measurement of success of emotional technology / interfaces
- Supporting user populations from specific Generations
- Supporting user populations with Physical Disabilities
- Supporting user populations with Intellectual Disabilities
- Creativity Support Systems
- Emotional Design issues / methods / experiences for novel interfaces including tangible, Mobile and Ubiquitous computing, mixed reality interfaces and multi-modal interfaces
- Usability
- User studies and fieldwork
- Methodological implications of Emotional User Studies
- Participatory design and Cooperative design techniques
- Ethical issues in emotional design
- HCI and Design education
- Eliciting User Requirements

The IADIS Interfaces and Human Computer Interaction 2011 conference received 155 submissions from more than 27 countries. Each submission has been anonymously reviewed by an average of five independent reviewers, to ensure that accepted submissions were of a high standard. Consequently only 37 full papers were approved which means an acceptance rate below 24 %. A few more papers were accepted as short papers, reflection papers, posters and doctoral consortia. An extended version of the best papers will be published in the IADIS International Journal on Computer Science and Information

Systems (ISSN: 1646-3692) and/or in the IADIS International Journal on WWW/Internet (ISSN: 1645-7641) and also in other selected journals including journals from Inderscience.

Besides the presentation of full papers, short papers, reflection papers, posters and doctoral consortia, the conference also included two keynote presentations from internationally distinguished researchers. We would therefore like to express our gratitude to Professor Andreas Holzinger, Head of Research Unit Human-Computer Interaction, Institute of Medical Informatics, Medical University Graz, Austria and Professor Harold Thimbleby, Professor of Computer Science at Swansea University, Wales, United Kingdom, for accepting our invitation as keynote speakers.

This volume has taken shape as a result of the contributions from a number of individuals. We are grateful to all authors who have submitted their papers to enrich the conference proceedings. We wish to thank all members of the organizing committee, delegates, invitees and guests whose contribution and involvement are crucial for the success of the conference.

Last but not the least, we hope that everybody will have a good time in Rome, and we invite all participants for the next year edition of the IADIS International Conference on Interfaces and Human Computer Interaction 2012, that will be held in Lisbon, Portugal.

Katherine Blashki,  
University of Sydney,  
Australia  
*Interfaces and Human Computer Interaction 2011 Conference Program Chair*

Piet Kommers, University of Twente, The Netherlands  
Pedro Isaías, Universidade Aberta (Portuguese Open University), Portugal  
*MCCSIS 2011 General Conference Co-Chairs*

Rome, Italy  
July 2011

# **PROGRAM COMMITTEE**

## **INTERFACES AND HUMAN COMPUTER INTERACTION CONFERENCE**

### **PROGRAM CHAIR**

Katherine Blashki, University of Sydney, Australia

### **MCCSIS GENERAL CONFERENCE CO-CHAIRS**

Piet Kommers, University of Twente, The Netherlands

Pedro Isaías, Universidade Aberta (Portuguese Open University), Portugal

## **INTERFACES AND HUMAN COMPUTER INTERACTION CONFERENCE**

### **COMMITTEE MEMBERS**

Aidan Rowe, University Of Alberta, Canada

Alessandro Casavola, University Of Calabria, Italy

Alexander Felfernig, Graz University Of Technology, Austria

Alexander Gelbukh, National Polytechnic Institute, Mexico

Alexander Kroener, Dfki Gmbh, Germany

Alexandros Potamianos, Technical University Of Crete, Greece

Amanda Meincke Melo, Unipampa - Federal University Of Pampa, Brazil

Ana Cristina Bicharra Garcia, Universidade Federal Fluminense, Brazil

Andrea Brogni, Istituto Italiano di Tecnologia, Italy

Andreas Auinger, University Of Applied Sciences Of Upper University, Austria

Angela Tinwell, Bolton University, United Kingdom

Ania Bobrowicz, School Of Engineering And Digital Arts, United Kingdom

Ann-Sofie Axelsson, Chalmers University Of Technology, Sweden

Antonio Maria Rinaldi, University Napoli Federico II, Italy

Aris Alissandrakis, University Of Hertfordshire, United Kingdom

Arttu Perttula, Tampere University Of Technology, Finland

Barbara Leporini, Istituto Di Scienza E Tecnologia Dell'informazione, Italy

Bente Meyer, University Of Aarhus, Denmark

Boris Brandherm, DFKI GmbH, Germany

Brunhild Bushoff, sagasnet, Germany

Carla Hoekendijk, Independant Consultant, Advisor Hogeschool Van Ams, Netherlands

Carsten Jessen, Danish School of Education, Aarhus University, Denmark

Caterina Senette, Istituto Di Informatica E Telematica, Cnr (pisa), Italy

Celmar Guimarães da Silva, Universidade Estadual De Campinas , Brazil

César Alberto Collazos Ordonez, University Of Cauca - Popayán, Colombia

Christopher Joyce, Anglia Ruskin University, United Kingdom

Clara Fernandez-Vara, Singapore-MIT GAMBIT Game Lab, USA



Connie Veugen, Vrije Universiteit Amsterdam, Netherlands  
 Corina Sas, Lancaster University , United Kingdom  
 Cristiano Maciel, University Federal De Mato Grosso, Brazil  
 Cyril Rebetez, SCEE, UK  
 Daniel Salber, mackey.nl, Netherlands  
 Darwin G. Caldwell, Italian Institute Of Technology (IIT), Italy  
 Elina Ollila, Nokia Research Center, Finland  
 Eric Fassbender, Charles Darwin University, Australia  
 Erik Champion, College Of Creative Arts, Massey University, New Zealand  
 Eugenio Aguirre, University Of Granada, Spain  
 Evan Golub, University Of Maryland, Usa  
 Francisco Jurado, Universidad De Castilla-la Mancha, Spain  
 Francisco Madera Ramirez, Universidy Autonoma De Yucatan, Mexico  
 Gerrit Meixner, German Research Center For Artificial Intelligence, Germany  
 Giulio Mori, Istituto Di Scienza E Tecnologie Dell'informazione, Italy  
 Grettel Barceló Alonso , National Polytechnic Institute, Mexico  
 Grigori Sidorov, National Polytechnic Institute, Mexico  
 Gunver Majgaard, University Of Southern Denmark, Denmark  
 Habib Fardoun, Instituto De Investigación En Informática De Albac, Spain  
 Hanna Wirman, University of the West of England, United Kingdom  
 Hassan Qudrat-Ullah, York University, Canada  
 Hideki Aoyama, Keio University, Japan  
 Jaakko Suominen, University Of Turku, Finland  
 Jaime Ramirez Castillo, University Of Castilla-la Mancha, Spain  
 Janne Yukiko Yoshikawa Oeiras Lachi, Federal University Of Grande Dourados, Brazil  
 Jean-Baptiste Labrune, Mit Medialab, USA  
 Jen Jenson, York University, Canada  
 Jesper Juul, New York University Game Center, USA  
 Joice Otsuka, Federal University Of São Carlos, Brazil  
 Jon Preston, Southern Polytechnic State University, USA  
 Jose De Castro, Ufc-virtual, Brazil  
 José Antonio Macías Iglesias, Universidad Autónoma De Madrid, Spain  
 José Luis Tapia Fabela, Universidad Autónoma Del Estado De México, Mexico  
 Julie Doyle, Dundalk Institute Of Technology, Ireland  
 Junia Anacleto, Federal University Of Sao Carlos, Brazil  
 Jussi Parikka, Anglia Ruskin University, Cambridge, United Kingdom  
 Kathia Oliveira, Université De Valenciennes Et Du Hainaut-cambrésis, France  
 Keith Nesbitt, University Of Newcastle, Australia  
 Ken Newman, University Of Applied Science Breda, Netherlands  
 Konstantin Mitgutsch, Massachusetts Institute of Technology, USA  
 Larbi Esmahi, Athabasca University, Canada  
 Leonardo Cunha de Miranda, University of Campinas (UNICAMP), Brazil  
 Leonardo Mattos, Italian Institute Of Technology, Italy  
 Leonie Ramondt, Anglia Ruskin University, United Kingdom  
 Liel Leibovitz, Columbia University, USA  
 Lisbeth Klastrup, IT University Of Copenhagen, Denmark  
 Luca Rossi, University Of Urbino, Italy

Luca Giulio Brayda, Istituto Italiano Di Tecnologia, Italy  
 Maki K. Habib, The American University In Cairo, Egypt  
 Manuel Cruz-hernandez, Immersion Canada Inc., Canada  
 Manuel Ortega Cantero , University Castilla - La Mancha, Spain  
 Marco Winckler, IRIT, France  
 Margarida Romero, Esade, University Ramon Llull, Spain  
 Maria Augusta Nunes, Universidade Federal De Sergipe (ufs), Brazil  
 Maria Claudia Buzzi , IIT - Cnr, Italy  
 Marina Buzzi, Istituto Di Informatica E Telematica, Cnr, Italy  
 Marjakankaanranta Kankaanranta, University Of Jyvaskyla, Finland  
 Mark Grimshaw, University Of Bolton, United Kingdom  
 Markus Wiemker, Rwth Aachen, Germany  
 Matjaz Gams, Jozef Stefan Institute, Slovenia  
 Mauro Dragone, University College Dublin (ucd), Ireland  
 Max Møller, Aarhus University, Denmark  
 Melissa Lee Price, Staffordshire University , United Kingdom  
 Michael Brown, Research, L1f2, Ordnance Survey , United Kingdom  
 Michael Hobbs, Deakin University, Australia  
 Michael Hohl, Sheffield Hallam University, United Kingdom  
 Michael O Grady, University College Dublin, Ireland  
 Michael Smyth, Edinburgh Napier University , United Kingdom  
 Michael Wagner, Kirchliche Pädagogische Hochschule Wien/Krems, Austria  
 Mirjam Eladhari, Gotland University, Sweden  
 Monica Bordegoni, Politecnico Di Milano, Italy  
 Natalie Ebenreuter, Alcatel-lucent Bell Labs France, France  
 Nicola Whitton, Manchester Metropolitan University, United Kingdom  
 Niki Lambropoulos, Intelligenesis, United Kingdom  
 Niklas Ravaja, Aalto University, Finland  
 Nitendra Rajput, IBM Research, India  
 Nuno Otero, University Of Minho, Portugal  
 Oscar Osvaldo Sandoval Gonzalez, Orizaba Institute Of Technology, Mexico  
 Otniel Portillo Rodriguez, Universidad Autónoma Del Estado De México, Mexico  
 Panagiotis Germanakos, Institute of Informatics and Telematics (IIT), CNR, Italy  
 Patricia Cristiane De Souza, Universidade Federal De Mato Grosso, Brazil  
 Patrick Felicia, Waterford Institute Of Technology, Ireland  
 Pauliina Tuomi, Tampere University Of Technology, Finland  
 Peder Burgaard, TDC, Denmark  
 Philip Breedon, Nottingham Trent University, United Kingdom  
 Pilar Lacasa, University Of Alcala, Spain  
 Ping Chen, University Of Houston-downtown, USA  
 Riad I. Hammoud, Delphi Electronics & Safety - World Headquarters, USA  
 Richard Bartle, University Of Essex, United Kingdom  
 Richard Hoadley, Anglia Ruskin University, United Kingdom  
 Robert Houghton, University of Nottingham, United Kingdom  
 Rodrigo Bonacin, Dssd - Cti, Brazil  
 Rommert Casimir, Tilburg University & Universitas Multimedia Nusantara,  
 Netherlands/Indonesia

Rosan Chow, UdK Berlin, Design Research Lab, Germany  
Ross Brown, Queensland University Of Technology, Australia  
Sarah Kettley, Nottingham Trent University, United Kingdom  
Sarah Sharples, University Of Nottingham, United Kingdom  
Sari Ostman, University Of Turku, Finland  
Sebastián Romero López, I.E.S Cañada De La Encina, Spain  
Seiji Isotani, Carnegie Mellon University, USA  
Stephen Barrass, University Of Canberra, Australia  
Sybille Caffiau, Lig/multicom, France  
Sybille Lammes, Utrecht University, Netherlands  
Thomas Markussen, Aarhus School Of Architecture, Denmark  
Timo Lainema, Turku School Of Economics, Finland  
Toni Granollers, University Of Lleida, Spain  
Troels Degn Johansson, Danish Design School, Denmark  
Vania Neris, Ufscar, Brazil  
Vicki Hanson, University Of Dundee, United Kingdom  
Victor M. R. Penichet, Escuela Superior De Ingeniería Informática (esii), Spain  
Weidong (tony) Huang, Csiro Ict Centre, Australia  
Weilian Su, Naval Postgraduate School - Monterey, USA  
Wouter Van Den Hoogen, Eindhoven University Of Technology, Netherlands  
Yulia Ledeneva, Autonomous University Of The State Of Mexico, Mexico

# **KEYNOTE LECTURES**

## **INTERACTING WITH INFORMATION: CHALLENGES IN HUMAN – COMPUTER INTERACTION AND INFORMATION RETRIEVAL**

**Professor Andreas Holzinger**  
**Head of Research Unit Human-Computer Interaction**  
**Institute of Medical Informatics, Medical University Graz, Austria**

### **ABSTRACT**

Today, we are confronted with a flood of information. Research in Human-Computer Interaction (HCI) and Information Retrieval (IR) has both long been working to develop methods that help people to identify, extract and understand useful information from data. The two fields, however, take very different perspectives in tackling the problem; and historically, they have had little collaboration. Let us for example look into the area of health: Medical professionals are faced with an increasing quantity of highly detailed, complex and non-standardized data at the press of a button, however, the time available to make decisions is the same as before the advent of such technological advances. According to (Gigerenzer, 2008) a typical medical doctor has approximately five minutes to make a decision. When everything turns out well no one complains; however, when something goes wrong, solicitors have nearly indefinite time to figure out whether and why a wrong decision has been made. The goal is to support medical professionals to interactively analyse information properties and visualize the most relevant parts without getting overwhelmed. The challenge is to bring HCI & IR to work together and hence reap the benefits, so that we can benefit medicine and health care even more.

# **INTERACTIVE NUMBERS - A GRAND CHALLENGE**

**Professor Harold Thimbleby**  
**Professor of Computer Science at Swansea University, Wales, United Kingdom**

## **ABSTRACT**

Numbers are used everywhere, but today numbers are mostly created and used interactively; they are not just passive written objects. People interact with numbers in almost every area of life. However, different styles of interactive number have different design trade-offs, particularly when human error is considered. To date, human error in interactive numbers has hardly been explored, resulting in most computer support for interactive numbers (from calculators to medical devices) being mediocre. Interactive number systems should be usable (effective, efficient, etc), free of idiosyncrasies and be demonstrably free from unacceptable levels of risk, particularly in environments where there are several different interactive number entry systems. We argue that defining “good” interactive number systems is indeed a serious challenge, but that substantial progress is being and can continue to be made, perhaps eventually resulting in an international standard for interactive numbers with solid empirical evidence of its value.

''''''''''''''''**Mg{ pqvg' Papers**



# INTERACTING WITH INFORMATION CHALLENGES IN HUMAN–COMPUTER INTERACTION AND INFORMATION RETRIEVAL

Andreas Holzinger

*Medical University Graz, Institute of Medical Informatics, Statistics and Documentation, Research Unit HCI  
Graz, Austria*

## ABSTRACT

Today, we are confronted with a flood of information. Research in Human-Computer Interaction (HCI) and Information Retrieval (IR) has both long been working to develop methods that help people to identify, extract and understand useful information from data. The two fields, however, take very different perspectives in tackling the problem; and historically, they have had little collaboration. Let us for example look into the area of health: Medical professionals are faced with an increasing quantity of highly detailed, complex and non-standardized data at the press of a button, however, the time available to make decisions is the same as before the advent of such technological advances. According to (Gigerenzer, 2008) a typical medical doctor has approximately five minutes to make a decision. When everything turns out well no one complains; however, when something goes wrong, solicitors have nearly indefinite time to figure out whether and why a wrong decision has been made. The goal is to support medical professionals to interactively analyse information properties and visualize the most relevant parts without getting overwhelmed. The challenge is to bring HCI & IR to work together and hence reap the benefits, so that we can benefit medicine and health care even more.

## KEYWORDS

Information Retrieval, Human-Computer Interaction, usability, complex data, complexity

## 1. INTRODUCTION AND MOTIVATION

Since the beginning of human existence, humankind has sought, organized and used information as it evolved patterns and practices of human information behaviour (Spink & Currier, 2006). Although the accessibility of information represents an important cultural advance, it also introduces a new challenge: retrieving only relevant information. To extract relevant information out of the vast complexity of data is the central quest of the modern information society.

However, accessing information is not always an easy task because we are dealing with the real world – where more data does not necessarily mean more information and more information is not always more knowledge. The challenge is that we have to consider the situation, the context. A further challenge is that mobile devices will be the primary tools in the future (Anderson & Rainie, 2008) and according to (Tsai et al., 2010) in mobile information retrieval there typically exist two main parts with their typical research fields: *Context Awareness* and *Content Adaption*.

Context Awareness deals with the fact that smart embedded devices have features to recognize the situation the user of the device is in at the moment. That means time, location, social status (social network).

Content Adaption mainly deals with how to present a user-friendly visualization of the results of information requests. Some examples on work for small screens can be found in (Noirhomme-Fraiture et al., 2005), (Sweeney & Crestani, 2004), (Jones, Buchanan & Thimbleby, 2002).

A further challenge is based on the fact that only a small percentage of data is structured – most of the data is semi-structured, weakly structured or even unstructured. A common misconception is to confuse structure with standardization. While the closely related fields of Information Retrieval and Knowledge Discovery have developed intelligent (semi)automatic processes and algorithms to extract useful knowledge from rapidly growing amounts of data, these methods fail when data are weakly structured and there is the



danger of modeling artifacts. Consequently, there are a lot of relevant research issues on the intersection of HCI and IR to help (medical) professionals to identify and extract useful information from data.

## **2. HUMAN INFORMATION BEHAVIOUR**

It is important to understand the four main terms that are essential when discussing human information behaviour (Spink & Saracevic, 1998), (Spink & Currier, 2006):

*Information Behaviour* is the totality of human behaviour in relation to sources and channels of information, including both active and passive information seeking, and information use. Thus, it includes face-to-face communication with others, as well as the passive reception of information as in, for example, watching TV advertisements, without any intention to act on the information given.

*Information Seeking Behaviour* is the purposive seeking for information as a consequence of a need to satisfy some goal. In the course of seeking, the individual may interact with manual information systems (such as a newspaper or a library), or with computer-based systems (such as the World Wide Web).

*Information Searching Behaviour* is the 'micro-level' of behaviour employed by the searcher in interacting with information systems of all kinds. It consists of all the interactions with the system, whether at the level of human computer interaction (for example, use of the mouse and clicks on links) or at the intellectual level (for example, adopting a Boolean search strategy or determining the criteria for deciding which of two books selected from adjacent places on a library shelf is most useful), which also involves mental acts, such as judging the relevance of data or information retrieved.

*Information Use Behaviour* consists of the physical and mental acts involved in incorporating the information found into the person's existing knowledge base. It may involve, therefore, physical acts such as marking sections in a text to note their importance or significance, as well as mental acts that involve, for example, comparison of new information with existing knowledge.

It is essential to understand the cognitive and perceptual abilities of the end users (Holzinger, Searle & Nischelwitzer, 2007). The best way is to understand the mental models of the respective end users. Mental models are defined as "cognitive representations of a problem [or information] situation or system" (Marchionini & Shneiderman, 1988), (Calero-Valdez et al., 2010). Some studies examined the role of users' mental model of an IR system in contributing towards search results and they argued that end users must have an appropriate mental model of a system in order to be able to use it to its full potential (Ahmed, McKnight & Oppenheim, 2004). Generally, the problem is that end users lack understanding of how the system operates. A good example is the study of (Dimitrioff, 1992), dealing with the relationship between the mental model of users and their search performance using a university search system. According to Dimitrioff, an accurate mental model of an online catalogue included eight components: the content of the database; the interactive nature of the system; the availability of more than one database; knowledge of multiple fields within records; knowledge of multiple indexes and/or inverted files; Boolean search capability; keyword search capability; and the use of a controlled vocabulary. The main barrier when studying mental models is the fact that they are not directly observable but must be studied by observing users' behaviour and is therefore hard to identify. Cognitive data such as users' knowledge, experience and expectations and how they cope with their information problem and interact with the system and interfaces are very important for the understanding of users' models of such systems.

## **3. USABILITY IN INFORMATION RETRIEVAL**

The first IR systems in the early 1970s allowed searches via command interfaces. The major disadvantage of such interfaces is the fact that the users must be familiar with the command language of the system to use it effectively (Hawkins, 1981), giving skilled experts the advantage.

Despite all improvements in end user interfaces, recent studies reported that web-based interfaces are still difficult to use and the need for better IR interface designs is still remaining (Ahmed, McKnight & Oppenheim, 2009). This work also emphasizes individual differences when using search engines and identified following main influence keys:

*Search Experience:* Describes the fact that users who are used to search engines are using a broader variety of the query language they depend on. Having search engine experience clearly has a positive impact und the user's search performance.

*Knowledge:* Knowledge about the topic enables the user to use more synonyms and combination of search terms to fulfill the information requirements. Generally Search Experience has to be presumed before the Knowledge factor takes effect.

*Academic Background:* Users with an academic background in science and engineering have a better search performance than people from the humanities given the same level of expertise. Interestingly experienced female searchers had better results than experienced male searchers.

*Users Age:* Older users with the same computer experience as younger ones had a lower success rate. On the other hand, a certain level of language understanding has to be assumed so older elementary students had better results than younger ones. Searchers with computational experiences overcome those with less computational background. Obviously computer experience influences the use of search engines in a positive way. Summarizing, all these factors influence the users in the searching strategy and experience in a positive or negative way. So there is a need for search engines that also take account the *personal context* of the user, in order to optimize their experience and success in fulfilling the information requirements. Consequently, an ideal system should make use of the abilities of both the human and the computer in tasks to which they are best suited, and to provide explanations of the data for enabling insights (Beale, 2007).

#### 4. COMPLEXITY AS MAIN CHALLENGE

Complexity is our main challenge, because most of our data is weakly structured or even unstructured and there is always the danger of modelling artefacts (which can then lead to wrong decisions). A good example are medical documents: The broad application of enterprise hospital information systems amasses large amounts of medical documents, which must be reviewed, observed and analyzed by human experts (Holzinger, Geierhofer & Errath, 2007). All essential documents of the patient records contain at least a certain portion of data which has been entered in non-standardized format (wrongly called *free-text*) and has long been in the focus of research (Gell, Oser & Schwarz, 1976). Although text can be *created* simple by the end-users, the support of automatic analysis is extremely difficult (Gregory, Mattison & Linde, 1995), (Holzinger et al., 2000), (Lovis, Baud & Planche, 2000). It is likely that some interesting and relevant relationships remain completely undiscovered, due to the fact that relevant data are scattered and no investigator has linked them together manually (Smalheiser & Swanson, 1998), (Holzinger et al., 2008).

Consequently, a major research area is on how to extract *knowledge* from this weakly or unstructured data. When we talk about structures, we will see some really interesting aspects of structures, in both microcosmos and macrocosmos.

A good example of a data intensive and highly complex microscopic structure is a yeast protein network. Yeasts are eukaryotic micro-organisms (fungi) with 1,500 currently known species, estimated to be only 1% of all yeast species. Yeasts are unicellular, typically measuring 4  $\mu\text{m}$  in diameter. The first protein interaction network was published by (Jeong et al., 2001). The problem with such structures is that they are very big and that there are so many. A great challenge is to find unknown structures (structural homologies, see e.g. (Jornvall et al., 1981)) amongst the enormous set of uncharacterized data. Let us illustrate this process with a typical example from the life sciences: X-ray crystallography is a standard method to analyse the arrangement of objects (atoms, molecules) within a crystal structure. This data contains the mean positions of the entities within the substance, their chemical relationship, and various others and the data is stored in a Protein Data Base (PDB). This database contains vast amounts of data. If a medical professional looks at the data, he or she sees only lengthy tables of numbers.

However, by application of a special visualization method, such structures can be made graphically visible and the medical professionals can understand these data more easily and most of all they can gain knowledge – for instance, it may lead to the discovery of new, unknown structures in order to modify drugs, and consequently to contribute to enhancing human health. The transformation of such information into knowledge is vital for the prevention and treatment of diseases (Wiltgen & Holzinger, 2005), (Wiltgen, Holzinger & Tilz, 2007).

To demonstrate that not only natural processes have such structures there is a nice example from (Hurst, 2007) which shows a visualization of the blogosphere (cf. also with (Leskovec et al., 2007)): The larger,

denser part of the blogosphere is characterized by socio-political discussion – the periphery contains some topical groupings. By showing only the links in the graph, we can get a far better look at the structure than if we include all the nodes.

A further example is from viral marketing. The idea is to spread indirect messages, which suggests spreading farther. If you press the *like*-button in Facebook – a similar process starts to an epidemic in medicine – a illness spreading through a population.

(Aral, 2011) calls it *behaviour contagion* and it is of much importance to know how behaviour can spread. We can mine masses of social network data in order to gain knowledge about the contagion of information. This is of particular interest for the health area.

## 5. CONCLUSION AND FUTURE OUTLOOK

Successful information retrieval systems will be those that bring the designer's model into harmony with the user's mental model. We can conclude that combining HCI with IR will provide benefits to our e-Society. Most of all, we must bridge Science and Engineering to answer fundamental questions on what is information and on (how) we can build such systems simply. Important future research aspects include:

- 1) research on the physics of information to contribute to fundamental research;
- 2) considering temporal and spatial information, in networks spatially distributed components raise fundamental issues on information exchange since available resources must be shared, allocated and re-used – Information is exchanged in both space AND time for decision making, therefore timeliness along with reliability and complexity constitute main issues and are most often ignored;
- 3) We still lack measures and meters to define and appraise the amount of information embodied in structure and organization – for example entropy of a structure;
- 4) considering information transfer: how can we assess, for example, the transfer of biological information;
- 5) Information and Knowledge: In many scientific contexts we are dealing only with data – without knowing precisely what these data represent. What is semantic information and how can we characterize it?
- 6) and most of all, we must gain value out of data – making data valuable.

Human-Computer Interaction and Information Retrieval (HCI & IR) is dedicated to contribute towards these challenges in their own ways; the challenge is to get them to do it collaboratively, and hence benefit medicine and health care even more.

## ACKNOWLEDGEMENTS

I am grateful for inspirations and help from Marcus Bloice, Sue Bogner, Philippe Palanque, Vimla Patel, Gig Searle, Klaus-Martin Simonic, Harold Thimbleby and Martina Ziefle.

## REFERENCES

- Ahmed, S. M. Z., McKnight, C. & Oppenheim, C. (2004) A study of users' performance and satisfaction with the Web of Science IR interface. *Journal of Information Science*, 30, 5, 459-68.
- Ahmed, S. M. Z., McKnight, C. & Oppenheim, C. (2009) A review of research on human-computer interfaces for online information retrieval systems. *Electronic Library*, 27, 1, 96-116.
- Anderson, J. Q. & Rainie, L. (2008), The Future of the Internet III. Online available: <http://www.pewinternet.org/Reports/2008/The-Future-of-the-Internet-III.aspx>, last access: 19.05.2011
- Aral, S. (2011) Identifying Social Influence: A Comment on Opinion Leadership and Social Contagion in New Product Diffusion. *Marketing Science*, 30, 2, 217-223.
- Beale, R. (2007) Supporting serendipity: Using ambient intelligence to augment user exploration for data mining and Web browsing. *International Journal of Human-Computer Studies*, 65, 5, 421-433.
- Calero-Valdez, A., Ziefle, M., Alagöz, F. & Holzinger, A. (2010) Mental Models of Menu Structures in Diabetes Assistants. In: Miesenberger, K., Klaus, J., Zagler, W., Karshmer, A. (Ed.) *Lecture Notes in Computer Science (LNCS 6180)*. Berlin, Heidelberg, New York, Springer, 584-591.

- Dimitrioff, A. (1992) Mental models theory and search outcome in a bibliographic retrieval system. *Library and Information Science Research*, 14, 2, 141-155.
- Gell, G., Oser, W. & Schwarz, G. (1976) Experiences with the AURA Free Text System. *Radiology*, 119, 105-109.
- Gigerenzer, G. (2008) *Gut Feelings: Short Cuts to Better Decision Making* London, Penguin.
- Gregory, J., Mattison, J. E. & Linde, C. (1995) Naming Notes - Transitions from Free-Text to Structured Entry. *Methods of Information in Medicine*, 34, 1-2, 57-67.
- Hawkins, D. T. (1981) Online information retrieval systems *Annual Review of Information Science and Technology*, 16, 171-208.
- Holzinger, A., Geierhofer, R. & Errath, M. (2007) Semantic Information in Medical Information Systems - from Data and Information to Knowledge: Facing Information Overload. In: Tochtermann, K., Haas, W., Kappe, F. & Scharl, A. (Eds.) *Proceedings of I-MEDIA '07 and I-SEMANTICS '07*. Graz, 323-330.
- Holzinger, A., Geierhofer, R., Modritscher, F. & Tatzl, R. (2008) Semantic Information in Medical Information Systems: Utilization of Text Mining Techniques to Analyze Medical Diagnoses. *Journal of Universal Computer Science*, 14, 22, 3781-3795.
- Holzinger, A., Kainz, A., Gell, G., Brunold, M. & Maurer, H. (2000). *Interactive Computer Assisted Formulation of Retrieval Requests for a Medical Information System using an Intelligent Tutoring System*. World Conference on Educational Multimedia, Hypermedia and Telecommunications ED-MEDIA 2000, Montreal, Charlottesville (VA): AACE, 431-436.
- Holzinger, A., Searle, G. & Nischelwitzer, A. (2007) On some Aspects of Improving Mobile Applications for the Elderly. In: Stephanidis, C. (Ed.) *Coping with Diversity in Universal Access, Research and Development Methods in Universal Access, Lecture Notes in Computer Science (LNCS 4554)*. Berlin, Heidelberg, New York, Springer, 923-932.
- Hurst, A. (2007), Data Mining: Text Mining, Visualization and Social Media. Online available: [http://datamining.typepad.com/data\\_mining/2007/01/the\\_blogosphere.html](http://datamining.typepad.com/data_mining/2007/01/the_blogosphere.html), last access: 2011-05-10
- Jeong, H., Mason, S. P., Barabasi, A. L. & Oltvai, Z. N. (2001) Lethality and centrality in protein networks. *Nature*, 411, 6833, 41-42.
- Jones, M., Buchanan, G. & Thimbleby, H. (2002) Sorting out searching on small screen devices. *Lecture Notes in Computer Science LNCS 2411*. Berlin, Heidelberg, New York, Springer, 81-94.
- Jornvall, H., Carlstrom, A., Pettersson, T., Jacobsson, B., Persson, M. & Mutt, V. (1981) Structural Homologies between Pre-albumin, gastrointestinal Pro-Hormones and other Proteins. *Nature*, 291, 5812, 261-263.
- Leskovec, J., McGlohon, M., Faloutsos, C., Glance, N. & Hurst, M. (2007) *Patterns of Cascading Behavior in Large Blog Graphs*. Philadelphia, Siam.
- Lovis, C., Baud, R. H. & Planche, P. (2000) Power of expression in the electronic patient record: structured data or narrative text? *International Journal of Medical Informatics*, 58, 101-110.
- Marchionini, G. & Shneiderman, B. (1988) Finding Facts versus Browsing Knowledge in Hypertext Systems. *IEEE Computer*, 21, 1, 70-80.
- Noirhomme-Fraiture, M., Randolet, F., Chittaro, L. & Custinne, G. (2005). *Data Visualizations on small and very small screens*. Proceedings of ASMDA 2005: 11th International Symposium on Applied Stochastic Models and Data Analysis, ENST, Brest (France).
- Smalheiser, N. R. & Swanson, D. R. (1998) Using ARROWSMITH: a computer-assisted approach to formulating and assessing scientific hypotheses. *Computer Methods and Programs in Biomedicine*, 57, 3, 149-153.
- Spink, A. & Currier, J. (2006) Towards an evolutionary perspective for human information behavior - An exploratory study. *Journal of Documentation*, 62, 2, 171-193.
- Spink, A. & Saracevic, T. (1998) Human-computer interaction in information retrieval: nature and manifestations of feedback. *Interacting with Computers*, 10, 3, 249-267.
- Sweeney, S. & Crestani, F. (2004) Supporting searching on small screen devices using summarisation. *Lecture notes in computer science LNCS 2954, Mobile HCI, Udine*. 187-201.
- Tsai, F. S., Etoh, M., Xie, X., Lee, W. C. & Yang, Q. (2010) Introduction to Mobile Information Retrieval. *Ieee Intelligent Systems*, 25, 1, 11-15.
- Wiltgen, M. & Holzinger, A. (2005) Visualization in Bioinformatics: Protein Structures with Physicochemical and Biological Annotations. In: Zara, J. & Sloup, J. (Eds.) *Central European Multimedia and Virtual Reality Conference (available in EG Eurographics Library)*. Prague, Czech Technical University (CTU), 69-74.
- Wiltgen, M., Holzinger, A. & Titz, G. P. (2007) Interactive Analysis and Visualization of Macromolecular Interfaces Between Proteins. In: Holzinger, A. (Ed.) *HCI and Usability for Medicine and Health Care. Third Symposium of the Workgroup Human-Computer Interaction and Usability Engineering (HCI&UE) of the Austrian Computer Society, USAB 2007, Lecture Notes in Computer Science (LNCS 4799)*. Berlin, Heidelberg, New York, Springer, 199-212.

# INTERACTIVE NUMBERS A GRAND CHALLENGE

Harold Thimbleby  
Swansea University  
Wales, SA2 8PP

## ABSTRACT

Numbers are used everywhere, but today numbers are mostly created and used interactively; they are not just passive written objects. People interact with numbers in almost every area of life. However, different styles of interactive number have different design trade-offs, particularly when human error is considered. To date, human error in interactive numbers has hardly been explored, resulting in most computer support for interactive numbers (from calculators to medical devices) being mediocre. Interactive number systems should be usable (effective, efficient, etc), free of idiosyncrasies and be demonstrably free from unacceptable levels of risk, particularly in environments where there are several different interactive number entry systems. We argue that defining “good” interactive number systems is indeed a serious challenge, but that substantial progress is being and can continue to be made, perhaps eventually resulting in an international standard for interactive numbers with solid empirical evidence of its value.

## KEYWORDS

Interactive numbers; user interface design; safety-critical user interfaces; medical systems; spreadsheets; calculators.

## 6. INTRODUCTION

Written around 1800 BC, the Rhind papyrus claims that

“accurate reckoning is the entrance into the knowledge of all existing things and all obscure secrets”

Indeed, numbers are the bedrock of civilisation. Mathematics has progressed since the earliest records of the denary number system, due to the Egyptians in 3100 BC. Notable steps to modernity include Leonardo Fibonacci of Pisa, who introduced Arabic notation (i.e., numbers written in modern Western digits, 0123456789) to the west with his *Liber Abaci* (The Book of Calculation, 1202); Gottlob Frege’s *Die Grundlagen der Arithmetik* (The Foundations of Arithmetic, 1884); then Kurt Gödel and Alan Turing’s arithmetisation of facts (1937), arguably the foundational result that allows computers to “do anything.” On the more practical side, Luca Paccioli’s *Summa de arithmetica, geometria, proportioni et proportionalita* (Everything about Arithmetic, Geometry, and Proportions, 1494) introduced double-entry book keeping, an idea that revolutionised accounting, and which lead directly to modern spreadsheets, via Dan Bricklin’s VisiCalc (1977) to today’s familiar Microsoft Excel (introduced in 1985).

History thus records a very long gestation of numbers as written objects (numerals) through to the modern concept of *interactive numbers*, a concept this paper introduces and explores. Compared to the widespread interest in written numbers and notations, to date very little attention has been paid to interactive numbers as such, with only a few exceptions, including the abacus, counting boards, slide rules and, in modern times, digital calculators, cash machines and of course computers. We shall show that interactive numbers are a surprisingly complex and rich area, and due to their ubiquity and use in high dependability applications, an area highly deserving of serious work: interactive numbers are a significant and promising research area. We contend that with more attention to interactive numbers, many everyday and critical systems could be even better.

Interactive numbers are created through interaction rather than just appearing in a fixed, written form; the new term is intended to focus attention on the non-trivial process of transferring a number from a human (perhaps though intermediaries and perhaps involving calculation) to a system.

With an interactive number we are explicitly concerned with how the user creates the number, considering the process as broken down into steps and possible errors and error corrections. For example:

- Using a spreadsheet, a user may key in the number 2.3. This is not something “that just happens” but involves selecting a cell in the spreadsheet, clicking, pressing a sequence of keys on the keyboard, then clicking elsewhere on the spreadsheet to complete the number entry. Each user action is *actively* handled by a computer program — unlike the simple case of writing a number on paper, where the paper does essentially nothing other than hold ink. Note that in a spreadsheet, a cell may not contain any number before the user starts to enter a number.
- Using a television (TV), the up/down keys and number keys can change the channel number. Typically pressing digit  $n$  will change to channel  $n$  (which in turn is mapped by the TV into a UHF frequency as defined during device set up). But if the user presses another digit,  $m$ , “quickly,” the channel will be  $10n+m$ , not  $m$ , or unless  $10n+m$  is greater than some small number (typically 30) the channel will be  $m$ . On some TVs, if  $n$  is pressed the TV will display “ $n-$ ” but not immediately change channel; if the user immediately keys  $m$ , the display will become “ $nm$ ,” that is, showing the Arabic numeral for the number  $10n+m$  (assuming  $10n+m < 30$ ) and the channel will change, or the user can wait, in which case — depending on the design — either the channel changes to  $n$  or the display “ $n-$ ” simply disappears and nothing further happens. These interaction rules are further complicated to account for the user keying up/down during number entry. Note that in a TV, unlike a spreadsheet cell, there is always a concept of “current number,” namely the current channel.
- In a hospital, a consultant may think of an appropriate drug dose for a patient and records this in an electronic patient record (EPR), hopefully without uncorrected errors. The pharmacy is informed, and prepares a drug. A nurse then programs an automatic delivery system (typically an infusion pump) to deliver the drug at the stated dose. The pharmacy may prepare the drug in standard sizes, meaning that the nurse will have to perform a calculation. Since there is a calculation involved, some hospitals will require two nurses to independently agree on the calculation; other hospitals think this is not as reliable as using just one nurse who therefore takes more care over the calculation as they are the only person responsible for the number. The calculated number is then entered in the drug delivery system, which will run Dose Error Reduction Software (DERS) to try to trap errors. Unfortunately, if the DERS, itself a complex sub-system, has been set up with the wrong drug, it may give a false sense of security. The patient may also have an active part in dosing, and be able to increase the dose on demand (perhaps for pain control), and a nurse may from time to time adjust the dose depending on the patient’s response to the drug. In short, then, a number (sometimes a formula) has been transmitted in a very complex way from the consultant’s head to the patient’s infusion device, with numerous opportunities for interaction — and error.
- Aviation accidents are frequently caused by problems with interactive numbers. The Air Inter Flight 148 in January 1992 performed a “controlled flight into terrain” killing 87 people in part because the autopilot (flight control unit, FCU) was probably in vertical speed mode instead of flight angle mode; the 33 the pilot entered was interpreted as 3,300 feet per minute descent not the intended 3.3 degree angle of descent (at the plane speed, equivalent to about 800 feet per minute). Following the crash, the FCU on A320s was modified to display vertical speed to four digits and angle to 2 digits, to make them less confusable.

In contrast to these interactive scenarios, it is conventional to think of numbers as static. Thus a written number, say 23.4, is viewed as a static representation of a number, and the process by which it was created or written down is of very little interest (except to school teachers). Ergonomics (human factors) has invested considerable research into readable numbers and good font design, notably helping develop OCR (optical character recognition) fonts that balance machine and human readability.

In some fields, notably medicine, great care is taken to ensure that numbers are written so that they are legible and unlikely to be misread.

Thus the Institute of Safe Medication Practices (ISMP, 2006) recommends that written numbers do not have “naked” decimal points (as in “.5” — because this may be misread as 5) and must not have trailing

zeros after the decimal point (as in “1.0” — because this may be misread as 10). Such rules help ensure that drug prescriptions are more dependable. The ISMP urge that “computer software vendors [should] make changes in electronic order entry programs” to follow their rules.

The ISMP and ISO rules are naturally considered as static rules: thus, under the ISMP rules, a written number is valid or invalid as an ISMP number. But what happens when a user keys in an interactive number, say 2.05? It is not obvious how and when to apply the rules: “2” is a valid ISMP number, “2.” isn’t, “2.0” isn’t, and finally “2.05” is. Worse, the ISMP rules are designed to reduce human error, but human error is unavoidable, and eventually users will create erroneous numbers. How should they be handled? Some interactive number systems display “0.” when they are switched on, and the display does not change if the user keys “.” (a decimal point); thus if the display shows “0.” the user cannot tell what pressing say “5” will do — the display will change to 0.5 or 5., depending on the mode the device is in. This is a simple example of the subtleties of interactive numbers.

When the user wants to enter 2.3 and that is what they do, the process may seem trivial. However, humans make slips, and occasionally a user intending to enter 2.3 will key in 3.2, 2..3 or some other sequence that is erroneous. What happens next is a question for interactive number entry. As it happens almost all interactive systems ignore the user’s error even when it is an obvious syntax error, like two decimal points (H.Thimbleby, 2010). Syntax errors are often ignored, treating the number as 0 or 2.3 as if nothing had happened. Some systems do use “data validation” (both syntactic and range checks) and may reject an error like 2..3 and hence prohibit the user from proceeding; nevertheless doing this is likely to increase the user’s stress and the likelihood of their making further errors! Unfortunately many systems do no validation (H.Thimbleby & Cairns, 2010).

To enter a number such as 2.3 the user employs an Arabic numeral keyboard, perhaps in telephone or calculator layout (and both layouts have minor variations). Which keyboard layout is better? In an environment that has both, is the “better” keyboard really better, or does it introduce more transfer errors? Or does variety encourage the user to pay more attention? Simply, we do not know.

We do know that users make slips, such as transposition errors (e.g., keying 2.3 as 23.). We know that very few interactive systems detect even obvious slips, such as two decimal points. We know that validating input to ISMP rules, would halve the out by ten error rate (where the entered number is ten times out from the intended number; this occurs with 2.3 entered as 23. — which happens to be a number with a naked decimal point). It is estimated that interactive numbers — not just handwriting — using ISMP rules would halve the fatality rate from miscalculations of drug doses (H.Thimbleby & Cairns, 2010).

Although they are fast, it is not obvious that Arabic numeral keyboards are the best way to enter interactive numbers — for some meaning of “best.”

Arabic numerals will be “best” for some applications (e.g., time paced copy typing) but we do not know the criteria because not enough research has been done. It may seem that Arabic numerals are obviously best for cases when numbers are used as names (examples being security codes, telephone numbers, credit card numbers), but the tradeoffs have not been explored. Even though we know that words are easier to remember, why are PIN codes numbers rather than words? It has, hitherto, seemed too obvious to need empirical research — and it would be expensive (because larger keyboards would be needed) to discover that numbers are suboptimal. Sadly, the cost of user error is often pushed on to the user, so until more research is done there is no incentive to improve the user interface. In high consequence, mission and safety critical areas, we think these widespread assumptions need reviewing and basing on solid empirical foundations.

One alternative is the “up/down” key approach, where pressing a “Λ” (or similar) key increases the number, and pressing a “V” key decreases the number. Like Arabic keyboards, there are many variations on this theme: for example, holding the Λ key down may cause the number to increase continually, and holding it down for longer may cause the number to increase faster. There may be several up and down keys, each with different rates (for example, changing different digits of the displayed number). Again, we do not know which styles are “better,” and if so for which applications.

Oladimeji *at al* (2011) have performed eye-tracking experiments on entering numbers. For Arabic style keyboards, 91% fixation time is on the keyboard and 9% on the display; in contrast for up/down interactive numbers it is 25% and 75% respectively. Users give much more attention to the display in an up/down system. Crucially, their unnoticed and uncorrected error rate is 6 times lower.

Users of up/down interactive number systems pay attention to the number they are entering, and moreover they have an “error correcting strategy” for normal use too: for example, to enter 95, they might quickly enter 100 then decrement it to 95. It seems clear that for any number entry where it is important that

the intended number is entered should be based on an up/down interface. Of course, unlike Arabic keys, up/down keyboards do not lend themselves to copy typing (typing while looking at text to be copied), and they may be slower. There is clearly a tradeoff, though 5 times the time to enter a number (6 seconds longer) compared to reducing the risk of death in a hospital (or an aircraft flying into the ground) might seem a good tradeoff to make.

Part of the time Oladimeji's experimental participants spent with the up/down interface was experimenting with it — making sure how it worked. In general, experimenting with an interface is a factor that would be expected to improve the reliability of interaction and (b) reduce with experience of consistent user interface design. Conversely, the lack of time spent experimenting with the deceptively “familiar” Arabic style of user interface means that users will be unfamiliar with the idiosyncratic ways these systems interact, especially how they handle errors. H.Thimbleby (2010) has shown an example where a user corrects an error to enter “5” but the system handled the user's error-correction in an unusual way and recorded “0.5” The device log therefore showed the incorrect value 0.5, so *obviously the user* made the error, as the log should have shown 5. We would argue that the device design (or a bug in the device implementation) led to the user being wrongly blamed for the incident.

All up/down interfaces we have explored have complicated features, such as time dependencies and hidden state. We have therefore designed a “rational” incremental up/down user interface, carefully following best design practice. Interestingly, to do this raises new design questions we do not know the answers for.

The science of human-computer interaction has ignored the devil in the details; it has assumed number entry is trivial and not worthy of attention. We have yet to prove that our design is formally correct, certainly an essential step towards improving dependable interactive numbers. In fact, it may be that the design is unique (up to certain trivial variations), and thus a *de facto* standard for improved interactive numbers.

We have also developed a principle-driven user interface for Arabic number entry; ironically, the current paper describing it was rejected by referees referring to it as merely “incremental” research! However, it is again noteworthy that a rigorously defined user interface even for something as “simple” as interactive numbers raises questions that there are no answers. More research is needed; conversely, it is increasingly apparent that all interactive number designs so far developed fall far short of the design standards that we should aspire to. Interactive numbers have been taken for granted.

See H.Thimbleby (2000) and W.Thimbleby & H.Thimbleby (2008) for wider discussion of number entry, calculators and mathematical user interfaces in general (one would be wrong in thinking they would obey mathematical laws).

W.Thimbleby's (2010) user interface uses a touch screen, and explores interactive numbers at a lower-level than usual, for instance allowing “ink editing” to edit a 1 into a 4 by overwriting the stroke  $\angle$ . In an experiment comparing a conventional Arabic keyboard calculator and his interactive number interface, W.Thimbleby (2010) had *no* errors, compared to about 60% for the conventional calculator.

## 7. SELECTED ISSUES IN INTERACTIVE NUMBERS

### 2.1 Bugs in Design or Problems in *any* design rules?

The ISMP (2006) rules should apply to interactive systems and they forbid naked decimals. Thus as the user keys “1.2” what do the rules say should be displayed after the decimal point has been pressed but before the 2 has been pressed? If the user enters an ISMP-invalid number such as “2.0” what should be displayed? As mentioned above, the rules assume that numbers are static and not interactive. Yet a system implementor has to make decisions at this level — what can be done?

Time outs are an obvious solution. After a time  $t$  of inactivity, the display will comply with ISMP rules. Unfortunately, for any time  $t$  selected, one can envisage cases where the time is too short or too long. Clearly a system should not reset if it interferes with the user; thus  $t$  should be longer than a typical interaction. Ten seconds? Yet a handover (e.g., one nurse replacing another) can be done in under 10 seconds, so the second nurse will experience the partially-completed interaction of the first, when it would have been preferable to have done the time out so the second nurse could start with a clean sheet. Many other scenarios are possible. However easy they are to implement, time outs do not solve the interaction problems; they are a proxy for



*knowing* when the user has lost track of the system mode. Proximity sensors or other techniques may be better. Reducing the hidden state of the user interface will also help.

In short it is clear that the simple ISMP rules are not adequate as they stand for interactive numbers. Programmers have to implement interactive numbers, and the rules (i.e., program) has to work at a finer level of detail than ISMP envisaged.

Either there is an interesting research programme to determine the best interactive number rules, or (possibly) no such rules exist that are completely consistent. In the meantime, developers will continue to implement suboptimal interactive number systems — there is no choice, as we do not know how to do it!

## 2.2 Numbers versus Numerals

It is usual to distinguish between numerals and numbers; the numeral being a concrete representation of the abstract number value. Confusion arises because we cannot speak about a number without representing it in some way. Thus 27 is an Arabic numeral that represents the number value 27, the same number value as  $26+1$  in fact. In Roman numerals, the value 27 would be represented as XXVII. As a binary numeral, it would be 11011, and as an English numeral “twenty seven.”

Since the user obviously creates a numeral, why do I use the terminology “interactive number” rather than “interactive numeral”? The reason is that the user is thinking of a number (which if you asked them, they might say in English or write down in Arabic numeral notation) and their intention is to get the number into the system they are using. They don’t want the system to have the numeral “twenty seven” they want it to have the value that is represented by twenty seven.

Of course many systems use Arabic notation for numerals, so the user will come to little harm if they think of a number as its Arabic numeral. However there are occasions when the differences are subtle and in some applications important. Thus, at first sight 027, 27 and 27.0, while different numerals, appear to represent the same number. Computer systems may think otherwise; Mathematica makes the difference  $27! - 27.0!$ , which you might expect to be zero, to be almost  $10^{14}$ ; or in Java  $27/4=6$  but  $27.0/4=6.75$ . In many programming languages, 027 (which seemingly has innocuous leading zeros) is 27 in octal, hence 23 in decimal, 4 less than the 27 programmer may have thought.

In some applications there may be no conventional representation of the number. A light dimmer would be an example: the user interacts with the dimmer to adjust the brightness of a light. The user perhaps does not think of this as an interactive number, but the user’s interaction adjusts the power going to the light. An electrician could connect a meter, and the meter would show the numerical value of the brightness. If the user so wished, they could remember this number and set the light to the exact same brightness on another occasion. In fact if the meter was permanently connected the user’s interaction would very obviously be interactive numbers. Some dimmers give the user a fixed number of choices, perhaps six: the user can interact to set the light strength anywhere from 0 (off) to 5 (full on).

## 2.3 Discrete versus Continuous Interaction

There is a considerable literature on number representation and processing (precision, resolution, stability, etc) in the numerical methods literature. We do not need to repeat it here except to note that it has concerned itself with number processing in a computer, without regard for the user. Much of the literature is relevant to interactive numbers however.

Interactive numbers can be *continuous*, taking on values that can, for the user, be imperceptibly close to each other, or *discrete*, taking on clearly distinct values. I’ve defined the terms from the user’s point of view; continuous numbers may be simulated on a digital computer by working to a high precision (e.g., fixed point, rational or floating point numbers) and conversely, real numbers may be simulated (e.g., on a calculator) by 8 digits of discrete numbers that are all clearly distinct to the user. Sometimes continuous numbers will be simulated by whole numbers; for example, the accelerator pedal in a car (if implemented by digital electronics) converts the angle of the sensor into an integer, which in turn controls the throttle and hence the RPM and speed of the car (which is what the user is concerned about); to the user, both RPM and speed will appear to be continuous even though they are controlled by small integers.

Sometimes continuous numbers are made to seem like discrete numbers. A car radio volume control may be a knob with detents that give the user the impression of discrete volume levels. Presumably the detents improve the user experience when turning the knob.

Users interact with discrete numbers by causing step changes, for instance pressing the key 7 typically multiplies the number by 10 and adds 7 (so if the display is 2, pressing 7 changes it to 27). Users interact with continuous numbers by turning knobs, moving sliders, holding buttons down for an interval while the number changes.

Users tend to come to a discrete number interface with a clear goal of the number they wish to enter. Users tend to come to a continuous number interface with a view to finding the best number by experimenting.

The behaviour of a user interface may depend on what has happened before. If everything relevant to the user is displayed (e.g., the number itself, perhaps plus a cursor) then the user interface is *declarative*. Otherwise, a discrete interface has *modes* and a continuous interface has *hysteresis*. Of course, one can simulate the other, and sometimes it may be clearer to partition the behaviour of a single interface into both modes and hysteresis.

The mathematical theories generally most appropriate for analysis of discrete and continuous systems are *graph theory* (including finite state systems) and *control theory*.

## 2.3 Low Uncorrected Error Rates, Low Error Rates

With interactive numbers, we are not so much interested in user errors, as in the errors that the user makes that do not get corrected: *uncorrected errors*. Most uncorrected errors are *unnoticed errors*. Noticed errors that are corrected will obviously slow the user down, so a design tradeoff is to balance low uncorrected error rate (which is good) with loss of speed (which is generally bad). It is of course possible for a non-error to be accidentally “corrected” and hence create an error, that itself may or may not be corrected; more likely is for an attempted error correction to fail to successfully correct a noticed error. Clearly eye fixation with the number display is likely to be essential for dependable corrections.

In some applications, correct numbers are more important than speed, whereas in some, the numbers being “good enough” (however that is defined) faster is more important. Interestingly, a user interface that gives the impression of having a very low error rate may encourage the user to miss errors when they do occur: a higher error rate may make the user more vigilant, and more skilled at error correction.

It is usual to distinguish between uncorrected user error and the magnitude of the harm that results; simply quantifying uncorrected errors is insufficient. If the numbers are monetary values, then typically any uncorrected error is unacceptable; whereas if the numbers are drug doses, then a relative error greater than some value (say 5%) would be unacceptable; other domains will have different criteria. A domain may additionally have *notifiable errors*, such as the number being out by a factor of ten or more.

User error rates during interactive number entry tests in the laboratory may be as low as 1–2% (Oladimeji *et al*, 2011), which hinders obtaining significant statistical results. If empirical tests compare different styles of interactive number entry (e.g., independently rotating digit dials versus dials that carry) the incident rates at the boundary cases will be even lower, perhaps 0.01% to 0.1%.

## 2.4 The Richness of Interactive Numbers...

Some interactive number systems are dedicated to particular domains (e.g., cash machines), and others are fairly generic (handheld calculators). Numbers also appear in many places where they are really names (e.g., telephone numbers, PIN numbers), for scales (e.g., the Richter scale), in historical contexts like “2/-” (“two shillings” or £0.1), and in contexts with “check digits,” as in International Standard Book Numbers (ISBNs), etc. Mobile phones handle telephone numbers, and therefore have to handle emergency numbers (999, 192, 911, etc), and — an interactive number issue — decide how to allow users to dial emergency numbers when a phone is otherwise disabled yet not dial it [often] by accident (e.g., from random knocks to buttons in a handbag). How numbers should be written is a complex subject too, itself not without controversy. Thus large numbers have thousands separators — unless they are page numbers or dates. And so on.

Most numbers are entered assuming that a dimensioned value is being entered. SI conventions on the use of k, M, n multipliers are clear, though there is conflict with ISMP guidelines (e.g., because a handwritten  $\mu$

is confusable with  $m$ ; and the solidus “/” may be misread as a 1; ISMP argues litres should be abbreviated L not l which would make 5 L look like 51; etc). Although most medical infusion pumps expect numbers in mL per hour, a nurse will have calculated the rate on a general purpose handheld calculator unaware of number dimensions, and hence unable to detect interactive dimension errors (H.Thimbleby, 2008).

A complete and consistent coverage of design issues for all domains is clearly impossible — calculators may be used for innovative interactive number uses beyond any reasonable design brief (e.g., many calculators can be set up so that pressing “+” makes them count events; this is a different sort of number entry approach than the usual 0–9 number entry envisaged by designers). Equally, a paper such as this is unable to cover all potentially relevant design issues: interactive numbers touch on many areas, and a complete review of issues is impossible — particularly in a short paper. However, the following topics deserve mention: number notations (e.g., scientific, SI, financial, ...); sets of preferred numbers that are within a specified relative error (e.g., the E6 preferred numbers, 10, 15, 22, 33, 47, 68 are within 20% of each other); and so forth.

Static number representations and fonts have been widely studied. Surprisingly, hard-to-read displays, such as seven-segment number displays, are widely used in safety critical applications, such as aircraft altimeters, personal radiation meters, etc (H.Thimbleby, 2011).

Although the SI/ISO 31-0/ISO/IEC 80000 standards specify that numbers “can be made more readable by separating them into groups, preferably groups of three, separated by a small space” few interactive systems do this. Some systems use the comma, even though this is specifically prohibited because of confusion with the decimal point (which is some countries is a comma). In other words, although there are standards for number notations, they are widely and regularly flouted. In the interactive case, the grouping changes as the user keys more digits: 1 234 becomes 12 345 when the user keys “5” — the space seems to move to the right, and its final position cannot be displayed correctly until the user has finished (or reached the decimal point).

*Numerical Notation* by Chrisomalis (2010) covers the history of static number notations in more detail than we need for this paper.

## 8. INTERACTIVE NUMBERS AS A GRAND CHALLENGE

Grand Challenges are big, worthwhile topics with recognized value to achieve. Typical Grand Challenges are to manage the nitrogen cycle, provide worldwide access to clean water, reverse engineer the brain, and more specialized but nonetheless strategic challenges like designing a verifiable programming language compiler. Well known Grand Challenges that have been achieved are to get man on the moon, to decode the human genome, and to eradicate smallpox. Grand Challenges unite a research community because they are worthwhile, and with concerted effort, achievable.

Interactive numbers is a “small” topic that in principle seems might, with a research programme, be completely specified, thoroughly evaluated, and moved into an international standard. Considerable benefits would ensue: every interactive number system would become easier and more reliable to use, and transfer errors would be reduced — presumably systems that looked the same would be the same. Yet as this paper showed, interactive numbers are deceptively simple, and a rigorous, empirically justified standard is a long way off. However, given the importance of numbers, particularly in safety- and mission-critical systems, a standard (no doubt with many subsections for various specialized domains) remains as a very worthwhile grand challenge.

Often Grand Challenges have spin offs. Thus the technologies that got man on the moon are much more widely used, for instance in launching satellites and in astronomy. Interactive numbers are a special case of HCI; they show that minute attention to detail has dividends. In researching interactive numbers, we will discover new fundamental ideas about interaction — numbers are just one form of state, and *all* interactive systems are about manipulating state.

The benefit of a Grand Challenge is focus on an achievable, worthwhile goal. Interactive numbers are ideal: unlike almost any other area of human computer interaction, they are focused; they have clear problems; the problems are soluble; and progress will benefit everyone. In contrast, the more “exciting” areas of research are often exciting precisely because they are so distracting and hard to pin down!

## 9. CONCLUSIONS

Since numbers are used everywhere in society, from gaming to international finance, from healthcare to nuclear power plants, from aircraft to climate change, from alarm clocks to GPS coordinates. Numbers are used in computer applications *known* to be buggy (spreadsheets, medical devices, etc), even small advances in the science and engineering of interactive numbers can be expected to make a major impact for good. It is time to take interactive numbers seriously. Moreover, our initial results, briefly discussed above, show that progress is certainly possible.

There is considerable scope to improve both Arabic number entry and incremental number entry; more software engineering research is needed; more empirical research is needed; and more application by programmers and developers to what is, in fact, a non-trivial but focused problem. In short HCI should take interactive numbers seriously. Amateurish *ad hoc* solutions — as almost all currently are, even on safety-critical systems, should no longer be acceptable, and we should be able to rigorously explain why they are unacceptable.

If it is to be called incremental research, then our field, HCI, indeed needs more incremental research and not just exciting novel research; every improvement in interactive numbers will benefit everyone (even a tiny increment multiplied up by millions is a big improvement), and benefitting people is why we pursue any research program in human-computer interaction anyway. Our goal, indeed a Grand Challenge, is a rigorous standard for interactive numbers, completely formally specified (with off the shelf definitive implementations), empirically justified in its effectiveness, and with guidance on design tradeoffs (e.g., for optimizing speed, accuracy, error rate, and manufacturing costs, etc).

## ACKNOWLEDGEMENTS

Paul Cairns, Abigail Cauchi, Andy Gimblett, Andreas Holzinger, Paolo Masci, Patrick Oladimeji and Will Thimbleby have all made essential contributions to the research summarised here. The research has been funded by EPSRC under grants EP/G059063/1 and EP/F020031/1.

## REFERENCES

- S. Chrisomalis (2010) *Numerical Notation*, Cambridge University Press.
- ISMP, Institute for Safe Medication Practices (2006) *List of error-prone abbreviations, symbols and dose designations*. See [www.ismp.org/tools/abbreviations](http://www.ismp.org/tools/abbreviations)
- P. Oladimeji, A. Cox, H. Thimbleby (2011), “Number Entry Interfaces and their Effects on Errors and Number Perception,” submitted.
- H. Thimbleby (2000) “Calculators are Needlessly Bad,” *International Journal of Human-Computer Studies*, 52(6): 1031–1069.
- H. Thimbleby (September+October, 2008) “Ignorance of interaction programming is killing people,” *ACM Interactions*, pp52–57.
- H. Thimbleby (2010) “Think! Interactive Systems Need Safety Locks,” *Journal of Computing and Information Technology*, 18(4):349–360.
- H. Thimbleby & P.Cairns (2010) “Reducing number entry errors: solving a widespread, serious problem,” *Journal of the Royal Society Interface*, 7(51), 1429–1439.
- H. Thimbleby (2011) “don’t use 7 segment displays,” *Proceedings BCS HCI 2011*, in press.
- W. Thimbleby (2010) *Drawing from Calculators*, PhD Thesis, Swansea University.
- W. Thimbleby & H. Thimbleby (2008) “Mathematical Mathematical User Interfaces,” in Proceedings Engineering Interactive Systems 2007/DSVIS 2007, *Lecture Notes in Computer Science*, 4940:519–535, Springer Verlag.



# **Full Papers**



# A COMPARISON OF IN-LAB AND SYNCHRONOUS REMOTE USABILITY TESTING METHODS: EFFECTIVENESS PERSPECTIVE

Mr. Obead Alhadreti<sup>1</sup>, Dr. Pam Mayhew<sup>1</sup> and Dr. Majed Alshamari<sup>2</sup>

<sup>1</sup>*University of East Anglia*

<sup>2</sup>*King Faisal University*

## ABSTRACT

Traditional in-lab usability testing has been used as the standard evaluation method for evaluating and improving the usability of software interfaces. However, in-lab testing, though effective, has its drawbacks, such as unavailability of representative end-users, high testing costs, and the difficulty of reproducing a user's everyday environment. To overcome these issues, various alternative usability evaluation methods (UEMs) have been developed over the past two decades. Among these, one of the most commonly used is the remote usability testing method. This paper is concerned with a comparative study of the traditional in-lab usability testing method and the synchronous remote usability testing method. It aims to examine how each method produces usability data, such as usability problems found, error number, time spent and success rate. The results of this paper discuss how these data differ based on the method used. It reveals some interesting results. Although the achieved data are similar, some measures differ significantly, such as identifying major usability problems.

## KEYWORDS

Remote usability testing, lab testing, synchronous remote testing.

## 1. INTRODUCTION

Usability is increasingly recognized as an important quality factor for interactive software systems in use today. Several studies have reported the benefits of a strong commitment to usability throughout the development life-cycle of a software product. Among the observable benefits of usable user interfaces, one can mention user productivity, performance, and safety and security. Usability is important not only in increasing the speed and accuracy of the range of tasks carried out by a range of users of a system, but also in ensuring the safety and security of that use (Jean, 2004). In order to determine the usability level of a software system, a number of different but related usability evaluation methods (UEMs) have been proposed over the last three decades. One of these evaluation methods is the traditional in-lab testing method, which has been used as the standard evaluation method for evaluating and improving the usability of software user interfaces. However, traditional in-lab testing, though effective, has its drawbacks such as the availability (or otherwise) of representative end-users, the high cost of testing, and lack of a true representation of a user's environment (Hartson, Jos *et al.*, 1998). To counteract these issues, various alternatives and less expensive usability evaluation methods have been developed over the past twenty years. One such UEM is the remote usability testing method. This method addresses the above issues by relying on real users conducting a number of real scenarios in their native environments. Remote usability testing is generally classified into synchronous (moderated) and asynchronous (unmoderated) testing (Brush, Morgan *et al.*, 2004). This paper is structured as follow: it begins with exploring the current usability evaluation methods (UEMs), including recent studies related to UEM comparisons. It then discusses the paper's objective and the approach taken. Data analysis and results are also discussed. It concludes with a brief discussion and conclusion.



## 2. RELATED WORK AND RESEARCH OBJECTIVE

Over the last three decades, numerous research studies have been carried out in attempts to overcome usability issues in software systems. As a result, many different types of evaluation methods have been developed (see Figure 1) (Jean, 2004). Of these methods, the most frequently used are usability testing methods, usability inspection methods, and model-based methods.

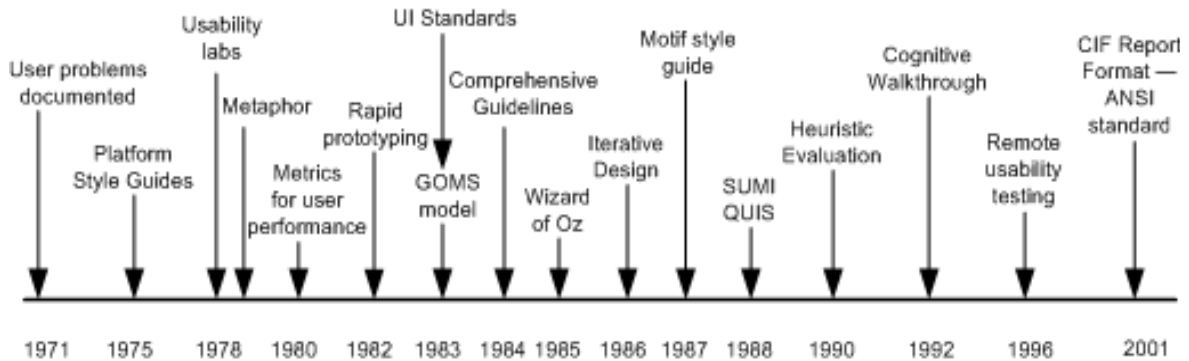


Figure 1. Development of usability evaluation methods (Jean, 2004)

Remote usability testing can be defined as, “usability evaluation where the test evaluators are separated in space and/or time from the test subjects” (Hartson, Jos *et al.*, 1996). The term “remote” refers mainly to the remote location of the test subject from the evaluator’s location (Castillo, 1997). As mentioned previously, remote evaluation can be generally classified into two main categories, synchronous remote usability testing (moderated), and asynchronous remote usability testing (unmoderated) (Susan and David, 2004). Synchronous remote evaluation, sometimes referred to as ‘live’ or ‘collaborative’ remote evaluation, is a usability evaluation method that has much in common with traditional in-lab usability evaluation (Selvaraj, 2004). It involves real users participating in the evaluation process from within their own environments using their own computers. In this evaluation method, the evaluator’s computer (in the usability lab) and the remote user’s computer (in their natural environment) are connected in real time through the Internet using a web-conferencing or screen-sharing application, and through an audio connection via the computer or a separate phone line. These allow the evaluator to collect data on the user’s actions by recording the test as the user performs the test tasks. The advantages of synchronous remote evaluation include the capability of collecting data from real users in their natural environment and the elimination of any need for participants to travel, also costs are lower making it more efficient, and more diverse users can be involved thereby including cultural contexts (Susan and David, 2004), (Morten Sieker, Henrik Villemann *et al.*, 2007). However, limited bandwidth and communication delays are some of the drawbacks of this method (Castillo, 1997).

The literature on remote synchronous testing method is more limited. Most of them conduct it by simulating in-lab settings (Susan and David, 2004), (Monty, Paul *et al.*, 1994). The rest adopted inspection methods. Tullis and others conducted research in order to evaluate the effectiveness of remote and lab usability testing methods (T. Tullis, Fleischman *et al.*, 2002). They reported the similar performance of these two methods. They found that remote testing users needed more time and completed more successful tasks than the lab group. They also received more negative feedback from those who attended lab testing (T. Tullis, Fleischman *et al.*, 2002). However, they did not report any statistical differences between the performances of these two methods. Andrzejczak and Dahai conducted a study that attempted to clarify the effect of testing location on usability test elements such as stress levels and user experience (Chris and Dahai, 2010). Although they reported that there are no differences between the performance of remote and lab testing, their study has a number of limitations such as users’ characteristics. All the users were students and the majority of them had previous experiences with the targeted website (Chris and Dahai, 2010). Another limitation is that the remote users were welcomed and briefed by the observer, which may encourage (or stress) the users to work hard in the test (Chris and Dahai, 2010). Their study shifted the focus to a new dimension in the comparisons of remote and lab testing; users’ experience and stress level. The literature lacks detailed experimental methods, data analysis and empirical data, although there are a number of important attempts such as (Brush, Morgan *et al.*, 2004), (McFadden, Hager *et al.*, 2002) and (Katherine, Evelyn *et al.*, 2004).

Although the literature usually compares two or more methods in terms of their efficiency, it should consider more whether or not the usability data are different if they are collected from different methods; remote synchronous usability testing and traditional in-lab usability testing. Such differences can mislead usability engineers who are required to produce a usability report at the end of their evaluation. Therefore, this paper examines whether or not the two methods, remote synchronous and traditional in-lab usability testing, produces different outputs. It questions the efficiency of these two methods.

### 3. METHOD

Two experiments were conducted in order to achieve the research objective. Two groups of users were involved. Each group consisted of 20 users to offer more validity to the comparison, as suggested by (Jacob Nielsen, 2006), (Brush, Morgan et al., 2004). This number can be also examined statistically. The users' gender, backgrounds, web experiences and ages were considered seriously in order not to influence the results later. Therefore, the characteristics of the users in each group were almost the same in terms of gender, age and Internet experience. These characteristics should reflect the targeted website audience, which is a social networking website. They all also perform the same tasks; each user performed five tasks. All the users were interested in and were familiar with the website; they participated for free. For the traditional in-lab testing section of this study, the participants carried out the test tasks in a usability lab at the university. Cam Studio screen capture software was also used to capture the participants' screen actions and record their voices on video (CamStudio, 2010). The video footages were then superimposed for further reviewing and analysis. The remote participants, on the other hand, were not provided with any equipment as they performed the test in their own environments. However, these participants were required to have a computer with Internet access, Internet Explorer 7 or higher, and a connected functional microphone. Skype Messenger 4.1 was used to connect and share the participants' screens with the researcher so that the participants' desktops could be observed for the duration of the test, and to communicate with the participants so that they could share their comments and suggestions (Skype, 2010). Although most test participants had Skype on their computers, participants who did not have Skype were advised to download it from the Internet. Before conducting a usability test, it is necessary to define clearly what metrics will be used to measure a system's usability level. According to (Sauro and Kindlund, 2005), the most frequently used usability metrics are: a) task completion rate; this concerns the percentage of tasks that are completed correctly during usability testing. Task completion rate provides a general picture of how the system being tested supports its end-users and the amount of improvement needed to make it work more effectively. b) Number of errors; this concerns the number of errors that participants make while performing the tasks. c) Time spent; this measures the time it takes a user to perform a single task from start to completion.

### 4. DATA ANALYSIS AND RESULTS

According to (Jacob Nielsen, 1995), usability problems can fall into one of the following categories of severity: not a usability problem, cosmetic, minor, major, and catastrophic. To ensure an objective assessment of the problems discovered in this study, the researcher sent the final set of problems identified by the two groups to a usability expert, who then classified their severity based on their frequency, impact and persistence. This classification has been used and recommended for use in usability testing in (Chen and Macredie, 2005). The following sections discuss how each of the two methods produces the usability data; number of problems discovered, task completion rate, time spent and error number. There are some similarities and differences between remote and in-lab testing. Interestingly, some of the differences were proven statistically.

Table 1 shows that both methods were able to reveal 41 usability problems. The in-lab group was slightly more effective than the remote group in identifying usability problems, as the in-lab group discovered (alone and with the remote group) almost 81% of the total number of problems, whereas the remote group only discovered (alone and with the in-lab group) almost 73% of the total number of problems. The in-lab group were able to reveal 27% of the problems discovered alone, whereas it failed to reveal 19% of the problems discovered only by the remote group. There were no statistical differences between the groups' performance

except in revealing major problems. The in-lab group performed better than the remote group in revealing major problems. The Fisher exact test reveals the significance ( $p = .01$ ). This means that there is 98.043% chance of the in-lab group revealing more major problems than the remote group. However, remote testing was more successful than in-lab testing in discovering minor problems. Generally, these results are in line with the findings of (Brush, Morgan et al., 2004), which concluded that in-lab testing outperformed synchronous remote testing in identifying usability problems, but they did not report any statistical differences (Brush, Morgan et al., 2004). These results are also in line with the (T. Tullis, Fleischman et al., 2002) study. They reported that lab users found more problems than remote users. In contrast, Sieker et al. reported the opposite; their results showed that synchronous remote testing outperformed in-lab testing in identifying unique major and catastrophic problems. Their findings were not sufficiently supported by statistics as the Fisher exact test did report any significant difference in the number of problems found by the two methods ( $p = 0.60$ ) (Morten Sieker, Henrik Villemann et al., 2007).

Table 1. Numbers and percentages of problems discovered

The used method	Cosmetic	Minor	Major	Catastrophic	Total
In-lab uniquely	2	2	5	2	11 (27%)
Remote uniquely	3	5	0	0	8 (19%)
Both	7	6	8	1	22 (54%)
Total	12	13	13	3	41 (100%)

Each participant was asked to perform five tasks on the targeted website, meaning that a total of 100 tasks were performed by each group. At the end of each task, the researcher assessed the completion rates and then classified them as successful (completed) or unsuccessful (not completed), as suggested by (Tom Tullis and Albert, 2008). Table 2 illustrates the results derived from the measurement of the task completion rate for both groups. The participants in the in-lab group successfully completed 63 tasks out of 100, whereas the participants in the remote group were only able to complete 57. These results support the findings of a study conducted by (Morten Sieker, Henrik Villemann et al., 2007), but are in contrast to the (T. Tullis, Fleischman et al., 2002) study. A possible explanation for this difference is that the psychological effect of the physical presence of the researcher and face-to-face communication with the in-lab participants; this might have given these participants more confidence when performing their tasks. This was reported in (J Nielsen, 2005), as users tend to work harder in the lab as they feel that they are under 'test' conditions, although they were informed that the website was the target of the 'test', not themselves.

Table 2. Task completion rate for each group

Tasks	Successful	Unsuccessful	Total
In-lab	63	37	100
Remote	57	43	100

Examining these results reveals that the participants in the in-lab group worked on the test tasks more quickly than the remote group. The in-lab group spent a total of 274 minutes and 8 seconds on the test tasks, whereas the remote group spent a total of 299 minutes and 5 seconds on them. The maximum time spent on a task was 5 minutes and 47 seconds, by one of the remote participants, while the minimum time spent on a task was 45 seconds, by one of the in-lab participants. This is in line with what has been reported: users work harder in the lab as they feel they want to do it correctly within an appropriate time (J Nielsen, 2005). This difference may be due to the equipment used in each testing method. The in-lab participants all used the same equipment under the same environmental conditions, whereas the remote participants used their own computers, which led to variations in the equipment used and may have affected the time required to perform the test tasks. These results are in agreement with the findings of (Morten Sieker, Henrik Villemann et al., 2007), (T. Tullis, Fleischman et al., 2002) and (Chris and Dahai, 2010), which show that in-lab testing participants complete tasks more quickly than remote testing participants. Figure 2 below compares the average time spent by each group on each individual task.

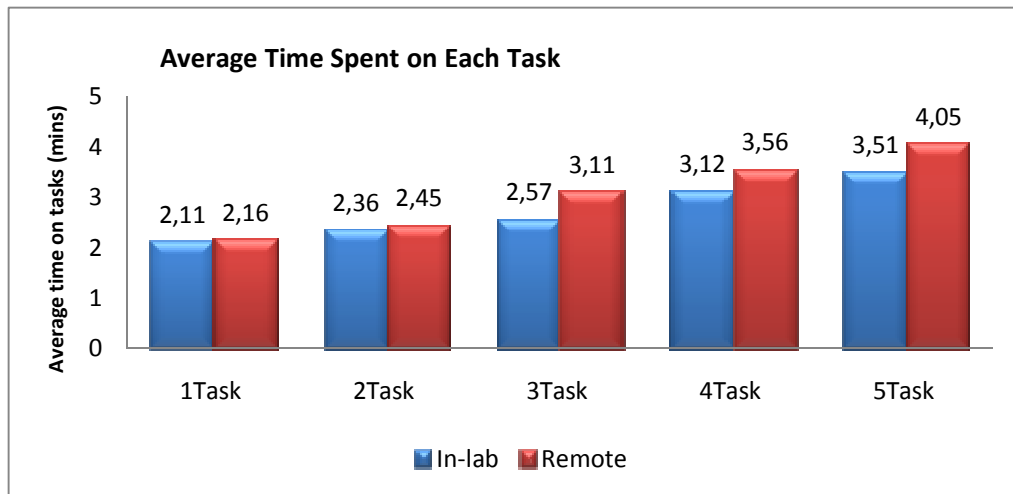


Figure 2. Average time spent on tasks

Table 3 shows the number of errors participants made on each task, the total number of errors on all tasks, and the average number of errors made by each participant.

Table 3. Number of errors on each task

UEMs	Task 1	Task 2	Task 3	Task 4	Task 5	Total	Average
In-lab	21	32	59	73	87	272	13.6
Remote	28	46	62	88	102	326	16.3

The total number of errors recorded for the remote group is larger than that recorded for the in-lab group. The difference between the in-lab and the remote groups' recorded errors might be due to the fact that, in the lab workstation, the researcher could ensure that the health, safety and ergonomic requirements for computer use were applied, whereas the researcher was unable to do the same for the remote group. Another reason might be that the in-lab participants were more concerned with how they would be judged by the researcher, who was located with them in the same place, hence they have tried to concentrate harder on the tasks. In general, these results support the findings of (Katherine, Evelyn et al., 2004), which concluded that remote testing participants make more errors than in-lab testing participants whilst performing tasks. The results in this section all suggest that the participants in the in-lab group were slightly more successful, efficient and accurate than the participants in the remote group, with regard to completing the test tasks.

In this research, it has been found that the users who spent more time, made more errors (in both groups), which can be seen clearly in Figure 3 below. This has been examined statistically and it was found that there is a strong statistical relationship between the time spent by users and the errors made by them ( $p = .001$ ). The  $p$  value is the same for each group and for the two groups together.

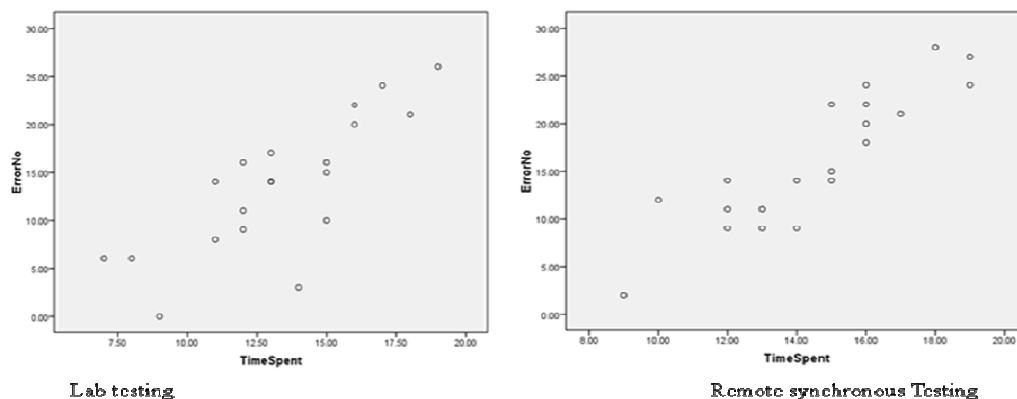


Figure 3. The strong correlation between time spent and errors made by users

## 5. DISCUSSION

Reliable study needs solid and clear steps. This study follows a systematic approach; one that is recommended for use in usability studies. All the usability factors (users number, characteristics, tasks number, the targeted website and usability measures) were taken into consideration in order to eliminate any influences that may occur or affect the achieved results. There are some interesting results; this study found significant differences between in-lab and remote usability testing in discovering major usability problems. No other differences were found between the two methods in terms of discovering catastrophic, minor and cosmetic usability problems. The results show that the two groups (i.e. methods) performed similarly, except that the in-lab group performed better than the remote group in discovering major problems; this was proved statistically. These results are in line with (Brush, Morgan et al., 2004), although they did not prove theirs statistically. The achieved results conflict with the (Morten Sieker, Henrik Villemann et al., 2007) study. The reasons behind this difference need further examination, such as investigating the most influential usability factors (test environment, tasks, user number, observer, the targeted website, users' characteristics and others).

The remote usability testing group needed more time, made more errors and performed fewer successful tasks than the in-lab group, and these results are in line with (Chris and Dahai, 2010), (Morten Sieker, Henrik Villemann et al., 2007) and (Katherine, Evelyn et al., 2004). There may be various reasons for the remote group's poorer performance, possibly related to those users' equipment such as their machines and Internet speeds. The other possible interpretation is that users tend to work harder in the lab than in normal circumstances (J Nielsen, 2005). If this is the case, the achieved results of the in-lab testing group may mislead usability engineers over the usability of the website. However, these performance differences (time, errors and success) were not proven statistically; further experiments are needed to clarify the reasons for these differences and to justify them.

## 6. CONCLUSIONS

In our comparison, we found some differences between in-lab and remote synchronous usability testing. We saw that there is a statistical difference between these two groups in terms of identifying major usability problems. The in-lab group revealed more major usability problems than the remote synchronous group. However, the in-lab group spent less time, made fewer errors and performed more successful tasks than the remote synchronous group. No statistical differences were reported for these differences. This research suggests that although the levels of efficiency for both methods are almost the same, other aspects should be investigated, including test cost, time and ease of application. Further investigation of these aspects, including method effectiveness, would be to the benefit of the website design sector.

## REFERENCES

- Brush, *et al.* (2004). A comparison of synchronous remote and local usability studies for an expert interface. CHI '04 extended abstracts on Human factors in computing systems. Vienna, Austria, ACM.
- Camstudio (2010). CamStudio. <http://camstudio.en.softonic.com>: Screen capture software.
- Castillo (1997). The User-Reported Critical Incident Method for Remote Usability Evaluation, Virginia Polytechnic Institute and State University. MSc.
- Chen and Macredie (2005). "The assessment of usability of electronic shopping: A heuristic evaluation." *International journal of Information Management* 25: 516-532.
- Chris and Dahai (2010). "The effect of testing location on usability testing performance, participant stress levels, and subjective testing experience." *J. Syst. Softw.* 83(7): 1258-1266.
- Hartson, *et al.* (1998). Remote evaluation for post-deployment usability improvement. Proceedings of the working conference on Advanced visual interfaces. L'Aquila, Italy, ACM.
- Hartson, *et al.* (1996). Remote evaluation: the network as an extension of the usability laboratory. Proceedings of the SIGCHI conference on Human factors in computing systems: common ground. Vancouver, British Columbia, Canada, ACM.

- Jean (2004). "Usability evaluation." Encyclopedia of Human-Computer Interaction IAD National Institute of Standards and Technology
- Katherine, *et al.* (2004). Here, there, anywhere: remote usability testing that works. Proceedings of the 5th conference on Information technology education. Salt Lake City, UT, USA, ACM.
- Mcfadden, *et al.* (2002). "Remote Usability Evaluation: Overview and Case Studies." International Journal of Human-Computer Interaction 14(3): 489 - 502.
- Monty, *et al.* (1994). "Remote usability testing." interactions 1(3): 21-25.
- Morten Sieker, *et al.* (2007). What happened to remote usability testing?: an empirical study of three methods. Proceedings of the SIGCHI conference on Human factors in computing systems. San Jose, California, USA, ACM.
- Nielsen (1995). Severity Ratings for Usability Problems. [www.useit.com](http://www.useit.com).
- Nielsen (2005). Authentic Behavior in User Testing. [www.useit.com](http://www.useit.com).
- Nielsen (2006). Quantitative Studies: How Many Users to Test? [www.useit.com](http://www.useit.com).
- Sauro and Kindlund (2005). A Method to Standardize Usability Metrics Into a Single Score. Conference on Human Factors in Computing Systems, Portland, Oregon, USA.
- Selvaraj (2004). Comparative Study of Synchronous Remote and Traditional In-Lab Usability Evaluation Methods, Virginia Polytechnic Institute and State University. MSc.
- Skype (2010). Skype. <http://www.skype.com/intl/en-gb/get-skype/on-your-computer/windows>: Screen capture software.
- Susan and David (2004). "Remote possibilities?: international usability testing at a distance." interactions 11(2): 10-17.
- Tullis and Albert (2008). Measuring the user experience. Burlington MA, Elsevier Inc.
- Tullis, *et al.* (2002). An Empirical Comparison of Lab and Remote Usability Testing of Web Sites. Usability Professionals Association Conference.

# A COMPARATIVE ANALYSIS OF THE MEANING OF 'LEARNABILITY' FOR CHILD AND ADULT USERS

Bester Chimbo, Helene Gelderblom and Ruth de Villiers  
*School of Computing, UNISA*  
*PO Box 392, Pretoria, 0003*

## ABSTRACT

The learnability principle relates to improving the usability of software, as well as users' performance and productivity. The principle was formulated mainly with adult users in mind and, although children are an important user group, fewer guidelines exist for their educational and entertainment applications. This study compares the groups, addressing the question: 'Does learnability of software interfaces have different meanings for children and adults?' A literature survey on learnability and learning processes, considered the meaning of learnability across generations. In an empirical study, users from 9–12 and from 35–50 were observed in a usability laboratory while learning to use educational software. Eye tracking data was also recorded. Insights emerged from the analysis, showing different tactics when children and adults use unfamiliar software, and revealing how they approach interfaces differently. Our re-interpretation of the learnability principle and the resulting design recommendations should help designers determine the varying needs of users of different ages, and improve the learnability of software designs.

## KEYWORDS

Learnability, Generational differences, Human-computer interaction, Child-computer interaction, Design guidelines, Usability.

## 1. INTRODUCTION

The basic principles and guidelines for software design were aimed at improving work performance and productivity – aspects relevant mainly to adult users (Pretorius, Gelderblom & Chimbo, 2010). Many of these principles are not relevant to children's products, since their needs, skills and expectations differ greatly (Chiasson & Gutwin, 2005). Children are a major user group and specific guidelines should be defined for their educational and edutainment software. Various human-computer interaction (HCI) researchers propose guidelines aimed at design for children (e.g., Fishel, 2001; Baumgarten, 2003; Gelderblom, 2008), and the present study attempts to augment these ventures.

Dix, Finlay, Abowd and Beale, (2004) suggest a comprehensive set of high-level directing principles to improve the usability of systems. They are divided into categories of learnability, flexibility and robustness. Learnability refers to the degree to which a user interface can be learnt quickly and effectively. This study investigates the learnability principle to establish whether it should be applied differently to software for varying age groups, with the goal of supporting software designers in meeting the needs of different generations. The results also contribute to the reformulation of the learnability principle in a way that distinguishes between adults and children.

The aim of our research was to compare the way in which children (aged 9 to 12) and adults (aged 35 and older) learn to use a new application. To this end, we set out to achieve the following:

- Identify interface aspects that are complex for adults, but not for children, and *vice versa*.
- Identify patterns in the learning behaviour of adults and children, respectively.
- Compare how children and adults learn to use software, thus extending the definition of the learnability principle.

A series of empirical studies was conducted with child and adult participants, using software aimed at children and software aimed at adults. Data was collected through interviews, observation, video recordings and eye tracking. This paper reports on studies that involved two software applications aimed at children. The

study involves descriptive research, as it portrays a picture of the details of a situation, social setting or relationship (Neuman, 2003). Rich descriptions of the behaviour of participants at the interface were the starting point for data analysis.

## **2. RELATED WORK**

### **2.1 The Learnability Principle**

Learnability (Nielsen, 1994; Shneiderman, 1997; Dix et al., 2004, Senapathi, 2005) is defined as the time it takes users to learn to use the commands for a task or the effort of a typical user in performing a set of tasks on an interactive system. Learnability comprises specific measurable attributes and can be evaluated by measuring these attributes in a real-life context. In the context of HCI, learnability relates to interactive features that help novice users learn quickly and that expedite steady progression to expertise.

Dix et al.'s (2004) classic set of principles are predictability, synthesizability, familiarity, generalisability and consistency. Predictability refers to the ease with which users determine the result of future interface interactions, based on the past interaction history. A predictable system is easy to learn. Synthesizability is the system's ability to provide an observable and informative notification about its internal changes of state (Aspinall, 2007). When a system is synthesizable, users can assess consequences of their actions. The familiarity principle relates to the ability of an interactive system to support the mapping of prior experiences onto a new system (Dix et al., 2004). Familiarity impacts on the novice's initial perception and whether he can determine the required actions from his own prior experience. Generalisability supports users in extending their knowledge of interaction in and across other applications, to new, but similar, situations. To support generalisability, consistency is essential, with systems offering similar functionality in comparable situations (Dix et al., 2004, Preece et al., 2007; Nielsen, 1994).

Although the definition of the learnability principle does not explicitly distinguish between adults and children, the literature on learning in general, provides evidence that children and adults learn differently.

### **2.2 Child and Adult Learning**

Learning is the process of transforming experience, skills and attitudes and involves various sub-processes. It is broader than education and can occur outside educational settings. Learning, whether deliberate or incidental, cognitive or practical, involves a change in knowledge, skills or attitudes. It occurs by using prior knowledge, conditions, and mental understanding to synthesize the skill or concept being acquired. Learning is flexible, occurring via different routes and learners do not always know exactly where they will end up (Rushton, et al., 2003; Goffree & Stroomberg, 1989; Jarvis, 2006).

Cotton (1995) distinguishes between three types of skills: psycho-motor skills that become automatic after repeated performance, perceptual skills controlled and conducted by the senses, and cognitive skills used in the process of acquiring knowledge. Playing on the computer is a composite skill that uses psycho-motor, perceptual and cognitive skills.

Learning in adults and children is qualitatively different due to the maturation of the brain's learning capacity and to different life situations (Illeris, 2006). Childhood learning is typically uncensored and trusting. Children develop their thinking abilities by interacting with other children, adults and the physical world. In adulthood, learning is fundamentally selective. Adults concentrate on learning things that relate to their work, careers, families and interests. Learning is motivated by a need to become more self-directed. Another important difference is that children have not reached the stage of complete cognitive development, emotionally or physically.

Adults are good at developing skills when the acquisition of these skills will add value to their lives. Children learn best when learning is initiated by their curiosity and interests, rather than imposed (Woolley, 1997). They naturally seek to make sense of experiences and to find order and patterns in their environment. They should know which behaviour produces desirable effects and eliminate those that do not (Ramey and Ramey, 2004).

Children are keen observers of other and tend to mimic their behaviour (Thornton, 2002). They learn by peer-interaction, playing together, or joint problem solving. Play gives opportunities for exploration and trial-



and-error in enjoyable, safe settings (Ramey & Ramey, 2004). When emotions are engaged, events and ideas are remembered better. Children learn well when involved in activities conducted by adults or more experienced children (Thornton, 2002). Certain differences between the learning of adults and children are particularly relevant to our research, namely:

1. Children often depend on adults for material and psychological support during learning, while adults depend upon themselves.

2. Adults are largely self-directed, in that they are responsible for deciding what, when, and how learning will occur. This difference is relative and varies according to context (Leberman et al., 2006).

3. Children perceive learning to be one of their major roles in life, while adults perceive themselves as doers who use learning to achieve success (Ference & Vockell, 1994). Adults learn best when they perceive the outcomes of the learning process as valuable and as contributions to their own development and success.

4. Adults have more life experience than children, which provides a good foundation for learning. However, it can also be a hindrance and a child's lesser experiences can occasionally prove more beneficial. Adults can be less willing to explore new ways of doing things.

The issue we now set out to investigate is whether and how these differences impact upon the learnability of software interfaces.

### 3. RESEARCH QUESTIONS AND METHODOLOGY

In this section we set out the main research question and subquestions; we describe the participants in the study; and explain how data was collected and analysed.

#### 3.1 Research Questions

The research was guided by the main question as to whether the learnability of software interfaces has a different meaning for children and adults. In the quest to solve this, the specific questions that we set out to answer were:

1. What differences in the learning behavior of adults and children can be observed when they learn to play an unfamiliar computer game?

2. Do these differences reflect the differences between adult and child learning that emerge from the literature on learning?

3. Should the learnability principle be interpreted differently when designing for adults and children?

4. What insights and design guidelines can be derived from comparing adult and child learning at the interface?

#### 3.2 Participants in the Study and the Software Employed

There were twenty-four volunteer participants in the usability laboratory studies — 12 children from 9 to 12, and 12 adults aged 35 to 50. The children attend primary school in Pretoria, South Africa and the adults are all academic or administrative employees at a university in Pretoria. All had moderately-high to high levels of computer literacy (See Table 1. In the experiments, some served as novices and others as experts. Five of the participants served both as experts and novices, as they were familiar with one of the games but not the other.

Table 1. Participant profile

	Age range	Gender	Computer game experience
Children	9 to 12	5 M, 7F	1 Low, 1 Moderate, 4 Moderately high, 6 High
Adults	35 to 50	6M, 6F	3 Low, 3 Moderate, 4 Moderately high, 2 High

The educational games used, were Timez Attack (Bigbrainz, 2005) and Storybook Weaver Deluxe 2004 (Broderbund, 2005). Timez Attack is an educational application, disguised as a captivating game to teach multiplication tables to children aged 7+. Users navigate an avatar through dungeons in search of golden keys to open doors. The keys take the form of multiplication sums. When a key is found, the program takes the

user through a sequence of events that systematically builds up the answer. Timez Attack uses the third person shooter genre of game play but, instead of weapons, answers to multiplication are used to defeat villains or open doors.

Storybook Weaver Deluxe 2004 is a software application for creating stories. Users choose from a variety of backgrounds to create scenes on the pages of their electronic story book and select from story characters and objects to create illustrations. The story is typed into the text panel. Users can add background music and sounds and, if a microphone is attached, they can record a voiceover for their story.

### 3.3 Data Collection and Analysis

A series of experiments was conducted in a formal usability lab, each involving a user learning a new application. The researchers observed different combinations of pairs of users (e.g. a child expert teaching an adult novice, a child expert teaching a child novice) as well as adult and child novices teaching themselves.

Data was collected through observation, eye tracking, and interviews. Observations were also video recorded for later analysis, which allowed repeated observation. With each iterative viewing, the researcher could change focus and note aspects she had not seen previously (Fetterman, 1998). Eye tracking is a technique for recording eye movement and eye-fixation patterns, and was used when single participants were teaching themselves how to use a software application. We also conducted informal, semi-structured interviews with both novice and expert participants after sessions, to discuss their reactions during the experiment. This clarified their thought processes, as they explained why they struggled at times.

We applied a five-step data analysis process proposed by Terre Blanche and Kelly (1999) that includes:

- *Familiarisation with and immersion* into the material gathered, reading (and watching) it repeatedly, to obtain thorough knowledge of the data;
- *Identification of themes*, which then formed the basis from which the descriptions of the observations could be refined and reorganized;
- *Coding*, to identify instances of specific themes, or relevance to specific themes;
- *Elaboration* to explore the newly organised material to identify similarities and differences in the data that may lead to new insights; and
- *Interpretation and checking* to ensure that there were no weak points.

The results provide insights into aspects of the learnability of software interfaces that adults and children approach differently.

## 4. RESULTS

The section is structured under the insights that emerged from the elaboration process. Insights are discussed along with evidence that justifies them, as well as with their implications for software design.

### **Insight 1: Children are more accepting of usability problems than adults.**

This was shown, for example, by reactions to an unexpected congratulatory message in Timez Attack. The message ‘Congratulations: Checkpoint reached’ incorrectly appears before the player has actually achieved any milestone. Children just ignored it! They waited for the message to disappear and continued with game play. Adults, however, were confused and questioned why they were being congratulated for doing nothing.

Further evidence comes from StoryBook Weaver in the form of different reactions by the two groups to the puzzling choice of selection icons. Children were undisturbed when they had to use the +, and not √, to accept the selected story object. They merely clicked on both buttons until they achieved the desired effect. Adults, on the other hand, were frustrated when they did not get the desired outcome by clicking on √. This shows that designers should not assume that a child’s impression of usability is valid. They can be very forgiving and may not comment on obvious problems. Products intended for children should be tested with adults to discover potential usability problems that may be overlooked when testing with children.

### **Insight 2: Adults tend to have fixed patterns as a result of life experiences and can be less open-minded during learning than children. Children, by contrast, learn in an *ad hoc* way.**

Children try out things just to see what happens. They do not expect anything to go wrong. Adults are more cautious and tend to be more self-critical than children. Adults are rigid in what they expect of a user

interface whereas children like to explore. In Timez Attack, snails have to be caught before the sums can be answered. The process of locating the snails follows an unpredictable sequence of mouse moves. When child novices played the game, they moved the mouse randomly, encountering snails and eliminating them. They asked questions that confirmed the above insight, for example: ‘Can I go the wrong way and see what happens?’ and ‘I want to see what happens if I do not catch one of the snails?’

While playing Timez Attack, children tried out different actions just to get ahead in the game. Eye tracking data show that they focused on the Play button that would activate the game, while adult participants fixated on the instructions at the bottom of the screen. Figures 4 and 5 show a child’s and an adult’s respective fixation patterns on the same screen. The adult has noticeably more fixations, indicating uncertainty.

In general, adults were more cautious in learning a new application. Cautiousness, as Salthouse (1991, p 176) points out, is ‘one of the most frequently mentioned performance-limiting factors’. Adult novices hesitated to make moves they were unsure about, as their questions show: ‘What if I click the wrong door?’, ‘How do I make the avatar walk to the left without making a mistake?’ and ‘What if I make a mistake and fall in a dungeon?’



Figure 4. Fixations of child novice on opening screen

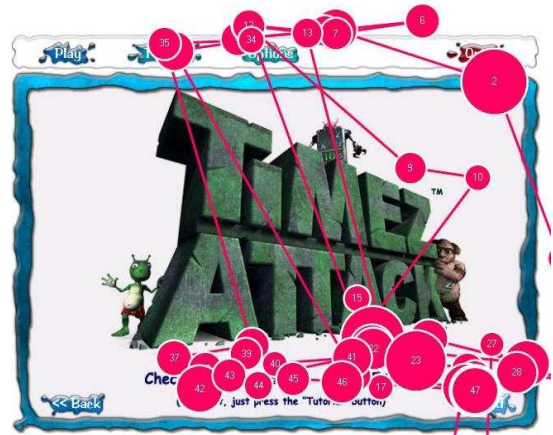


Figure 5. Fixations of adult novice on opening screen

Applications designed for children should show greater tolerance for incorrect operation than applications for adults. Designers may work through a risk assessment to ensure that the applications and their implementations do not expose children to unacceptable risks. If an application requires a very specific sequence of actions, it should be made clear so that users do not follow a wrong trail.

**Insight 3: Child novices can be faster than adult novices in mastering mouse and keyboard navigation skills during game play.**

In Timez Attack, all the adults struggled to use the mouse/keyboard combination to walk the avatar through the dungeons. Examples of the emotional expressions of adult novices in reaction to mouse/keyboard navigation difficulties included: ‘Oh man, what is the mouse doing?’ and ‘Good grief! What must I do with this thing?’

Children displayed greater dexterity than adults when using the mouse to navigate. Of the seven child novices who learnt how to play Timez Attack, only one experienced difficulty with navigation, and that only initially.

If software is designed for all age groups, they should provide customizable user interfaces that enable users to choose the user interface controls to match their preferences. Moreover, designers should not assume that, because adults generally have better hand-eye coordination, they will be better than (or just as good as) children at navigation.

**Insight 4: Adults want to have a clear and holistic picture of the entire software application before they start using it, whereas children just start using the application.**

This insight is evidenced by how adult experts taught the novices. They would begin by asking the novice if they knew anything about the application and would explain what the software was about, before giving instructions for using it and demonstrating it practically. Child experts immediately instructed novices on how to use the software applications and allowed the novices to participate from the start.

Child novices rarely asked for assistance during game play. They discovered things by themselves while adult novices asked for help as soon as they were given the chance to play the game. Adults were hesitant to try anything independently and asked many questions.

All the adult participants began by reading the tutorials while only one child participant read the Timez Attack tutorial. This confirms that adults want a clear and complete picture of what is coming.

Designers should provide appropriate guided tours of the application for first-time adult users. Since children prefer to get on with it, designers should not rely on them using the tutorials. If both children and adults will use an application, designers should provide customisable guided tours to accommodate both types of users.

**Insight 5: Children often accept what they are learning regardless of its purpose. Adults find learning to be purposeful if it has meaning and adds value to their lives.**

Children were clearly more engaged, and were not put off easily by usability issues and other problems. The fact that adults showed signs of frustration much earlier and often struggled where children did not, can be linked to the fact that the applications were not ones they would use voluntarily. Children and adults get more engaged and involved if they are using software that relates to them personally. Adults connect their learning of new software applications to life experiences that may include work-related activities, family responsibilities, and even previous educational experiences. In Storybook Weaver, child novices produced pictures related to their fantasy worlds, whilst adult novices produced mature pictures related to their day-to-day work or social environment.

## 5. DISCUSSION

The main research question of this study relates to the possibility that learnability may have different meanings for users of different ages. We consider this in the discussion that follows on the subprinciples of learnability (defined in Section 2.2) relating it to the insights gained through this study (Section 4). We also discuss what we have learnt in terms of the learning process and end with some recommendations for design that emerged from the results.

### 5.1 Reinterpretation of the Subprinciples of Learnability

The literature review established that a system would be easy to learn if it was predictable. Predictability allows users to know beforehand what will happen when they click on a menu item or press a key. Insight 1 showed, for example, that children did not consider the meanings of the two buttons in Figure 2 as critically as the adults did. They merely tried them out until one worked, while adults showed some confusion when the buttons did not function as they would have predicted. Elements relating to predictability were also evident from Insight 2. Children used trial-and-error to play the game, whilst adults relied on instructions. Eye tracking results showed that children's fixations were longest on the Play button whilst the fixations of adults were longest on the instructions.

A user interface that adheres to the principle of synthesizability allows the user to understand which user actions have led to the current state, what the system did to get there, and what the user should expect next. Insight 2 relates to synthesizability. It refers to the broader life experiences of adult novices that allow them to develop mental models that may enable them to overcome the difficulties of learning to use a new application. However, at other times, these mental models may be detrimental to learning. Adults' fixed patterns can cause them to be less open-minded than children to new learning. Children use whatever they learn through trial-and-error to construct cognitive maps of the workings of an unfamiliar software application.

Familiarity is the degree to which the user's own real-world personal experience and knowledge can be drawn upon to derive insights into the workings of an unfamiliar system. When the system has familiar elements, the user will relate it to similar, real-world situations or systems, thereby reducing the amount of cognitive burden to become adept at using it. Insight 2 also relates to familiarity.

A system is generalisable if users are able to use what they have already learnt to conduct new tasks. Insight 3 relates to generalisability, suggesting that child novices mastered mouse and keyboard navigation skills faster than adult novices. Besides youthful dexterity, as opposed to a general slowdown of motor co-

ordination with age, the main source of performance advantage in mouse and keyboard mastery by child novices over adult novices was found in the generalisability of mouse and keyboard skills mastered in other prior applications. Consistency applies when the system behaves in the same way when comparable sequences of actions occur in similar situations. Consistent interfaces are easier to learn and use (Preece et al., 2007). They help users gain confidence in using systems, and encourage them to try exploratory learning strategies (Nielsen, 1994). Insights 2 and 4 support Nielsen's assertion that consistent user interfaces encourage exploratory learning strategies. If system feedback and responses are consistent, child novices will have more success in their exploratory approach to learning.

The findings of this study suggest different interpretations of the learnability principle and its subprinciples for children and adults. To summarise:

Predictability is more crucial in adult products than in those aimed at 9 to 12 year olds. Adults need to be sure of what happens next and what they are allowed to do next, while children are more willing to just explore. Children are less concerned about the effects of their actions than adults. Synthesizability has a different meaning for the two user groups since their differing levels of experience will influence the way they form mental models about the working of a system. Designers should thus be aware that children may construct different mental models from what they (the adult designers) would expect.

A system that adheres to the principle of familiarity for adults may include elements with which children are not familiar. On the other hand, the fact that many children are exposed to technology from an early age, may mean that new input mechanisms, to which they have been exposed via computer games, may be unfamiliar to older people. The consequences of generalisability and consistency on learnability may be different for each user group, but our findings did not show that the two user groups understood the meanings of the two principles differently.

Not all the insights could be related to the existing subprinciples of learnability. We therefore identified the need to incorporate the concept of 'engagement' into the definition of learnability, since the users' level of engagement can determine their commitment to learn the application. This will be a topic for further research into the re-definition of the learnability principle.

## 5.2 The Learning Process

Literature on the learning process alludes to differences in the way that adults and children learn. Insight 2 highlights the differences in the life experiences of adults and children as the source of observed differences between them.

Insight 5 relates to the theory of social constructivism, which emphasizes the importance of the learner being actively and personally involved in the learning process (de Villiers, 2005). Children are simply happy to accept what they are learning regardless of its purpose, but adults appreciate learning if it brings added value and meaning to their lives.

Literature on the characteristics of adult learners, states that adult learners need to know why they should learn something before undertaking to learn it. Insight 4 relates to this characteristic when it refers to children trying out new things just to see what happens and not being worried that something might go wrong. Adults, on the other hand are more cautious, and tend to be more self-critical than children. Adults are rigid in what they expect of a user interface, whereas children like to explore.

Our results thus show that the differences in how adults and children learn in general, do not always apply when they are learning to use a new software application:

- Whereas adults usually rely on themselves in the learning process, when they learned to use the software games in this study, they depended more on the support of instructions and outside help than children did.

- In these games adults were not noticeably more self-directed in their learning than children.

- Adults' broader life experience did not have a clear effect on their learning of the games. In some instances it hindered them, rather than helped them. Children were more confident to learn through trial-and-error, while adults preferred to read instructions.

These findings on the differences between the general learning styles of children and adults could serve as the basis for recommendations to application designers to better satisfy their intended end users.

### 5.3 Recommendations for Design

Based on the results of this study, the following recommendations summarise the implications for design:

1. Software intended for children should not rely on written instructions. Software for adults, on the other hand, should provide detailed instructions and Help facilities.
2. If adults will use the product, tutorials that give product overviews should be provided.
3. Applications designed for children should show greater tolerance for incorrect operation.
4. When designers produce software applications intended for adults, they should make the value of using it clearly apparent. Given the rationale for learning something, adults will be more likely to invest time in it.
5. The different user groups have their own skills and abilities, therefore design principles for one group may not be applicable to the other group. Designers should acknowledge that they, as adults, may not understand the needs of child users. On the other hand, when designing for children, they should test the usability of their designs with adults, because children are more tolerant of usability problems.

## 6. CONCLUSION

The lack of sources in the literature that deal specifically with the learnability principle, suggests a gap in the body of knowledge. This study was an attempt at filling that gap, but more work needs to be done to improve the granularity in the description of the subprinciples. Specifically, better distinction between some of the principles, for example, generalisability and consistency, could be achieved through further research. The importance of 'engageability' to be incorporated into the definition of learnability has been identified as worthy of further investigation.

A limitation of the study was that the experiments were conducted in a usability lab that isolated participants in a controlled environment, where they could interact only with the facilitator and complete tasks with only the tools provided for them. This may have introduced bias into the results of the study. Users' behaviour can be influenced by the fact that they are being observed.

External validity, or the generalisability of the study, is limited by the fact that there were only 24 participants. Although the children represented different cultural groups and home languages, they all attended two schools situated in relatively privileged areas. The adult participants were from the same workplace, but they did represent a range of skill levels (from a full professor to a security guard) and cultural groups.

The study highlighted the need for software designers to distinguish actively between child and adult users. The results lead to reformulation and re-interpretation of the learnability subprinciples to differentiate between the needs of adults and children and served as foundation for specific recommendations for design. We believe that our findings will help to advance HCI practice and to improve the quality of software targeted at different age groups.

## REFERENCES

- Aspinall, D. 2007. *Principles and Guidelines interface Design Rules: Informatics*. UK: University of Edinburgh.
- Baumgarten, M. 2003. Kids and the Internet: A Developmental Summary. *ACM Computers in Entertainment*, 1, Article 2.
- BigBrainz, 2005. *TimezAttack*. BigBrainz, Inc.
- Broderbund, 2005. *Storybook Weaver Deluxe 2004*. Riverdeep Interactive Learning Limited
- Chiasson, S. and Gutwin, C. 2005. Design Principles for Children's Software, *Technical Report. HCI-TR-05-02*, Computer Science Department, University of Saskatchewan.
- Cotton, J. 1995. *The theory of learning: An introduction*. Kogan Page Limited, London.
- De Villiers, M. R. 2005. e-Learning artefacts: are they based on learning theory? *Alternation* 12.1b: pp.345-371.
- Dix, A., Finlay, J., Abowd, G. and Beal, R. 2004. *Human-Computer Interaction*. Pearson Education.
- Driscoll, M. 2000. *Psychology of learning for instruction*. 2nd ed. Needham Heights, MA: Allyn and Bacon.

- Ference, P. and Vockell, E. 1994. Adult learning characteristics and effective software instruction. *Educational Technology*, 34:23-25.
- Fetterman, D. M. 1998. *Ethnography: Step by step*. 2nd ed. Thousand Oaks, CA: Sage.
- Fishel, C. 2001. *Designing for Children*. Rockport Publishers, Inc.
- Gelderblom, J. H. 2008. Designing technology for young children: *Guidelines grounded in a literature investigation on child development and children's technology*. PhD Thesis, University of South Africa.
- Goffree, F. and Stroomberg, H. (eds.) 1989. *Creating adult learning, theoretical and practical communications on educational design for adults*. SMD, Leyden.
- Illeris, K. 2006. *How I learn: learning and non-learning in school and beyond*. Roskilde University Press.
- Jarvis, P. 2006. *Towards a comprehensive theory of human learning*. Routledge, London.
- Leberman, S., McDonald, L. and Doyle, S. 2006. *The transfer of learning. Participants' perspectives of adult education and training*. Gower Publishing Ltd, England.
- Neuman, W. L. 2003. *Social Research Methods - Qualitative and Quantitative Approaches*. 5<sup>th</sup> ed. USA: Pearson Education.
- Nielsen, J. 1994. *Usability engineering*. Morgan Kaufmann San Francisco.
- Preece, J., Rogers, Y. & Sharp, H. 2007. *Interaction design: Beyond human-computer interaction*. 2nd ed. John Wiley & Sons.
- Pretorius, M., Gelderblom, J.H. and Chimbo, B. 2010. Using Eye-Tracking to compare how adults and children learn to use an unfamiliar computer game. In: Kotze, P., Gerber, A., van der Merwe, A. & Bidwell, N.(Eds). *Proceedings of SAICSIT 2010 ACM International Conference Proceedings Series*. Bela Bela, October 2010.
- Ramey, C. T. and Ramey, S. L. 2004. *How children learn and how parents can help: Center for Development and Learning*. Covington, LA 70433.
- Rushton, S., Eitelgeorge, J. and Zickafoose, R. 2003. Connecting Brian Cambourne's conditions of learning theory to brain/mind principles: Implications for early childhood educators. *Early Childhood Education Journal*, 31(1).
- Salthouse, T. A. 1991. *Theoretical perspectives on cognitive aging*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Senapathi, M. 2005. A Framework for the Evaluation of CASE Tool: Learnability in Educational Environments. *Journal of Information Technology Education*, Volume 4.
- Siemens, G. 2006. 'Connectivism: Learning theory or pastime for the self-amused.' Available from: <<http://davidtjones.wordpress.com/2009/10/04/learning-theories-and-e-learning>> [Accessed: 18 June 2010].
- Terre Blanche, M. and Kelly, K. 1999. *Interpretive Methods*. In Terre Blanche, M. and Durrheim, K (Eds) *Research in practice: Applied methods for the social sciences*. Cape Town, UCT Press.
- Thornton, S. 2002. *Growing minds*. Harvard University Press, London.
- Venezky, R. and Osin, L. 1991. *The Intelligent Design of Computer-Assisted Instruction*. New York: Longman.
- Woolley, J. 1997. Thinking about fantasy: Are children fundamentally different thinkers and believers from adults? *Child Development*. 68(6) (1997), 991-1011.

# A SEMIOTIC PERSPECTIVE TO WEB USABILITY: AN EMPIRICAL CASE STUDY

Muhammad Nazrul Islam

*Turku Centre for Computer Science (TUCS), Department of Information Technologies  
Åbo Akademi University, Turku, Finland*

## ABSTRACT

In the present day User Interfaces (UIs) are complicated software components, which play a crucial role in the usability of web applications. An explosion on interface design for HCI has been commenced over the last decade. But very little attention has been paid to semiotics theories for web interface design, though designing the web sign has a widely acceptable crucial effect on enhancing users understanding and satisfaction. For these, the objective of this paper is to reflect user experiences in interface signs interpretation and how these could affect the usability of web applications. To accomplish this objective, a systematic empirical case study was conducted on a web application. This study was replicated with seven participants from five different educational institutions in Finland and followed a strict case study methodology to ensure the validity and reliability of our research outcomes.

## KEYWORDS

Think aloud usability testing, semiotics, interface signs, web application, user experiences

## 1. INTRODUCTION

Over the past years, with the advent of globalization and the rise of information technology, it has become obvious that one of the most important qualities of web application would be the ease by which the end user can learn and interact with these applications. Thus, the activity of assessing the quality degree of the applications is becoming an arduous task. Users' degree of satisfaction in using as well as interacting with a web application established the quality of this application (Triacca, 2003). The most significant measurement unit of satisfaction is usability, as it is "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" (ISO 9241-11, 1998). Thus, usability of a system contributes to user happiness, satisfaction, as well as pleasure; and conversely a lack of this contributes to user dissatisfaction and frustration, and thus eventually will result in the total abandonment of the system. Therefore, usability is considered a key quality for a web application.

The web interface plays the main role for the interaction between human and computer in web applications. The growing demand of the present Internet world leads us to focus on designing these web interfaces, user perception on web interface signs, web usability as well as the crucial roles of designing web interface signs to HCI. In fact, these design principles are semiotics by nature and semiotics is the science of signs (Peirce, 1932-58), that is, a theory about sense production and interpretation.

Therefore, the purpose of this research was to discuss semiotics theories and to show the significance of semiotics to design and also to evaluate the web interface signs to boost web usability. Indeed, this research shows that semiotics is one of the important fundamental design dimensions that affect the usability of a web application. In this research a systematic empirical case study on a web application was conducted and revealed how user's interpretation to interpret interface signs could affect the overall usability concerns. The complete study report with all data sets can be found in (Islam, 2011).

This paper is structured as follows. In section 2, the previous research related to our work is described. Semiotics theories and these relations to interface sign interpretation are discussed in section 3. These semiotics theories were also the sources of motivation to design and articulate this paradigm. The steps of experimental method for empirical study are discussed in section 4. An empirical case study on a web application, the Ovi calendar of Ovi by Nokia (<http://calendar.ovi.com>) is discussed in section 5. In section 6,



our research findings are presented. Finally, the conclusion concerning semiotics perspective to interface sign design as well as ideas of future research is provided in section 7.

## 2. STATE OF THE ART AND RELATED WORKS

Over the last few decades, usability evaluation method has been considered as an important quality assessment technique in website evaluations (Al-wabi & Al\_khalif, 2009). Thus, different Usability Evaluation Methods (UEM) have emerged and been developed in research and practice in the field of usability. UEM can be divided into four classes: analytic methods, specialist reports (usability inspection), observational methods (usability testing) and user reports (survey) (Whitefield, A. et al, 1991). Analytic methods are mainly driven by analysis (Diaper & Stanton, 2003) of tasks that need to be done by the end users. Usability inspections focus on feedback from experts in HCI or web application design. Heuristic evaluation (Nielsen, 1999) (Nielsen & Mack, 1994), cognitive walkthroughs (Hertzman & Jacobsen, 2003), feature inspections (Nielsen & Mack, 1994) are more common evaluation methods to this group. Checklists, usability principals or rules are used as guidelines to direct this kind of evaluations. Co-discovery (Barnum & Dragga, 2001), think-aloud (Hertzman & Jacobsen, 2003) (Nielsen, 1993) are effective example methods of usability testing group. In co-discovery, two or more users work together in the evaluation. For the think-aloud, a small number of users are involved individually, users verbalize while using the system to complete the given task to express his/her thoughts, feelings, and opinions. The final group, user report involves the use of questionnaires and interviews for data collection (Nielsen, 1993) (Usability Net, 2010).

All the evaluation methodologies presented above are lacking the evaluation of semiotic issues of web applications. These methods do not analyze the intrinsic values of user interface, specially the interface signs of user interface. To allow the analysis of intrinsic values of interface signs during usability evaluation, a semiotic engineering approach has been evolved (Souise, 2005). However, current well-structured web usability evaluation methods and techniques consider semiotic aspects as generic criteria for evaluating the user satisfaction, often confusing and blending them with other usability problems (i.e. problems related to content or to layout design) (Triacca, 2003). Moreover, very few methods give the right importance of semiotic design and evaluation to optimize the web usability.

The main reasons to skip semiotics issues in the currently available UEM of web applications as well as for designing the interface signs are: (i) lack of knowledge on semiotics and its theories in general, (ii) lack of theoretical background on semiotics theories to web interface sign design and its evaluation, (iii) lack of understanding the necessities of semiotics to interface design and evaluation, (iv) lack of awareness on how semiotically designed interface signs affect the web usability, etc. This research has mainly focused on these issues and shows how users' understanding of interface signs affects web usability, and thus eventually presents the significance of semiotics theories to design and evaluate the interface signs through a systematic empirical case study on a web application.

## 3. SEMIOTICS AND SIGN INTERPRETATION

Signs take the form of words, images, sounds, odors, flavors, acts or objects, but these things have no intrinsic as well as intended meaning and these things become signs only when designers provide these with meaning (or, sense) (Morris, 1938). Few examples of Ovi calendar interface signs are presented in figure 1. Among the many different semiotics models two models are presented here which were more relevant to this research work: (i) Peirce's semiotics model (Peirce, 1932-58) consists of a triadic relationship containing: the *representamen* (*representation or sign*) - this stands to somebody for something in some respect or capacity. It addresses somebody and creates in the mind of that person an equivalent, or perhaps more developed sign; the *object* (*referent*) - is the actual thing the sign stands for and the *interpretant* (*meaning*) - is therefore the sign created in the mind of the perceiver or the reaction caused by the object in the perceiver (Andersen, 1992). For these, a sign requires the concurrent presence of these three constituents. (ii) Semiotics theory by Gottlob Frege's terms for the three vertices of the semiotic triangle were *Zeichen* (sign) for the symbol, *Sinn* (sense) for the concept, and *Bedeutung* (reference) for the object (Frege, 1879). As an example of the semiotics triangle, Frege cited the terms '*morning star*' and '*evening star*' and both terms refer to the planet

Venus as their meaning, but their senses are very unlike the way in which the planet is presented (one term refers to a star seen in the morning, and other one refers to a star seen in the evening). Therefore, there is no one-to-one link between the object and the sign; various signs may have a single meaning in spite of several meanings. Different signs vehicles can refer to the same object since each sign vehicle has its own flavor or sense that leads it to the same object. Therefore, users generally guess the sign meanings through the *creation and interpretation* of 'signs'.


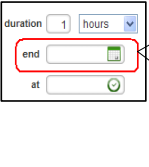



Sign	Interpretant	Object	Interpretation
	Designer P4	Retrieve deleted item	Accurate
	Designer P1	Textbox accept input data (date) in fixed format (dd / mm / yyyy) from keyboard or by cursor interacting with the calendar icon append with text box. Textbox accept input data (date) only from keyboard (do not get any hints about acceptable date input format) and appended calendar icon has no interactivity but it used only to give an indication that this text box is for date value.	Moderate
	Designer P2	Create / add a new category Create / add a new event	Conflict
	Designer P7	See all deleted items See all stored / entered items	Erroneous
	Designer P7	Choose option to select whether changes (of entered events with repetition) will effect to all repeated events or only the current one.	Incapable

Figure 1. Examples of participants' interpretation of interface signs and its categorization

For these, the users' interpretations (few examples are depicted in figure 1) of interface signs were classified into the following categories based on the accuracy level of user interpretation with respect to the designer's interpretation for an interface sign: *a) accurate*- user's interpretation completely matches the designer's interpretation and this category reflects the semiotics theory, *(b) moderate*- user's felt more than one distinct object, one of which was the right one about the interface signs and probability to obtain the right object at the first attempt may be less than the accurate interpretation (for example, if a user proceeds with a sign to obtain a particular object but the sign does not really stands for that), *(c) conflicting*- user's felt more than one distinct object in his/her mind about the interface signs and user felt confused about choosing the right object that will match to the designers intention, *(d) erroneous*- user's interpretation referred to a completely different object other than the designer's interpretation, and *(e) incapable*- user could not able to interpret the interface sign at all. These categorizations were also used in empirical studies in section 5.

#### 4. EXPERIMENTAL METHOD

In this research, two user tests were conducted (i) Interface sign intuitive test and (ii) Conventional think-aloud usability test to reach at research goal. To obtain the research outcome through these tests the six sequential steps were followed. Briefly these are the six steps. *Step 1*: The problem statement and test objectives were clearly defined to reflect the purposes of conducting the tests and appropriately derive the remaining steps. *Step 2*: Tasks list were prepared and then all the interface signs were listed along with listing (separately) the entire related interface signs (heuristically) to these tasks for the web application being

tested. *Step 3*: Participants who might be the user of this studied application were recruited and scheduled. *Step 4*: A user test were conducted to understand the user interpretations of these listed interface signs and collected the data in a systematic way. This test is named here as interface sign intuitive test. *Step 5*: User testing to perform the given tasks was conducted following the conventional laboratory based think-aloud method and collected the data in a systematic way. *Step 6*: Finally, these tests data were analyzed and examined to observe the user behavior focusing the users' understanding of interface signs and how these understandings affect users' performance. For example, an interface sign S is related to a task T. From the interface sign intuitive test, if it was happened that a user U does not understand properly the intended meaning of S. Then, usability testing data were examined to observe: (a) how the user U behaves to perform the task T while the sign is S related to this task T was not properly understandable to U. (b) how these behaviors influence the web usability. The studies ended by discussing the important observations that emerged from the analysis phase, and also presented the future trends of semiotics theory as applied to interface design and evaluation.

## 5. AN EMPIRICAL STUDY ON WEB APPLICATION

The purpose of our empirical study following the above systematic procedure was to observe the user experiences to deal with interface signs and how these could affect the web usability. This study was conducted on an example web application, the Ovi calendar of Ovi by Nokia during the late 2010 at usability testing laboratory of Åbo Akademi University, Finland. This section briefly discusses how the empirical study was conducted to reach at our research goal.

### 5.1 Problem Statement and Test Objective

This study mainly focuses on user understanding of interface signs of the web application from a semiotics perspective. One basic research question is addressed: *How do interface signs (semiotics) affect web usability?* This study objective was to obtain the answer of this research question. In particular, this research wanted to observe:

- user understanding (accuracy level) in interpreting the intended meaning of interface signs.
- user behavior to perform a specific task with respect to his/her understanding (accuracy level) of the task-related interface signs.
- And how their behavior influenced web usability.

### 5.2 Tasks and Interface Signs

A set of scenarios were created where each scenario contained multiple tasks. The scenarios were written in the language of user's tasks. The scenarios and its related tasks are briefly presented in table 1. After finalizing the tasks list, an inspection (heuristically) was carried out very meticulously to find (i) all interface signs, and (ii) related interface signs to each task. Then two lists of interface signs were prepared: one having all interface signs of the Ovi calendar and another having related interface signs to each task.

### 5.3 Participants

Anyone who wants a personal, free calendar service that can be accessed from any location from any web browser might be the users of this product. Therefore, students were chosen as our test users. Due to limitations of time and money, this study did not cover other types of users. A series of questionnaires were designed to qualify the potential users. The overall study involved seven participants aged 21 - 30, selected from five different universities in Finland. All participants had good experience in using the personal computer, the internet, the real world calendar and three users had prior experience in using a web calendar, but no participant had prior experience in using the Ovi calendar (see table 2).

Table 1. List of scenarios and related tasks

Scenarios	Task no.	Tasks
<i>Log-in &amp; event entry</i>	1	log-in to Ovi calendar (data was provided)
	2	create an event
	3	create an event with advanced options
<i>Search &amp; edit event</i>	4	search for an event (event entered previously)
	5	edit an event
<i>check, delete and log-out</i>	6	check weekly event list
	7	delete an event
	8	log-out to leave Ovi calendar

Table 2. Test participants profile in brief (H: High, M: Medium, L:Low, N: None)

Features	P1	P2	P3	P4	P5	P6	P7
Familiarity with personal computer	H	H	H	H	H	H	H
Internet familiarity	H	H	H	H	H	H	H
Age	21	22	22	26	25	29	24
Education	M	M	M	H	M	H	H
Familiarity with real world calendar	H	H	H	H	H	H	H
Familiarity with online Ovi calendar	N	N	N	N	N	N	N
Familiarity with other online calendar	L	N	H	N	N	H	N

## 5.4 Interface Sign Intuitive Test

The interface sign intuitive test was conducted through user interviewing mainly. The main reasons for choosing interviewing were (Online Resource, 2010): interviews need very few facilities, easy to organize, enjoyable, as well as a good way to find in-depth information about users. The user interview was conducted one by one following the thinking aloud method (Lewis, 1982). At the beginning, the interviewer gave a very short lecture to the participant regarding the purpose of the interview. The questionnaires used to conduct this test session were: *What do you think about the intended meaning of this sign? / What could be the purpose of using this sign? / What is your guess about the referential content for this sign? / Why does this sign stand for?* The author as interviewer and a participant, as interviewee were seated together in front of a computer, showed the list of all interface signs of web application being tested. Selected participants were asked to formally interpret these interface signs (base on the questionnaires raised for this test session), talking aloud and described their understanding of each interface signs. Page snapshots of studied application from where signs were listed were also showed and they were asked to “re-comment” on any signs if they thought their past comments were not appropriate to any particular sign. The fundamental purpose of asking these questions was to obtain an indication of their understanding and classify their interpretation into: accurate, moderate, conflicting, erroneous and incapable (see section 3). The interviewer noted these data during test sessions and these entries were checked again with the video record of the test sessions.

## 5.5 Think-aloud Usability Test

To perform a usability test with each participant, the following activities were followed. A short lecture was given about the usability testing in general. Activities during test sessions consisted of observing users performing their tasks in a usability test laboratory. Their activities were recorded in videos and they were observed through a one way mirror. Post-task questionnaires were used to obtain immediate feedback of the users after completing each scenario. The users were asked about the ease and difficulty of tasks and provided options to write comments on different issues they felt during the completion of tasks. It helped to obtain feedback when users' memory was fresh. At the end, when users finished last scenario, post-test questionnaires were delivered. Later the video record of the test sessions were examined and coded using data-logging software (Observer 5.0) to obtain test data. Apart from this, more data were collected from different kinds of questionnaires, for example pre-test, post-task, and post-test.

## 5.6 Analysis and Examination of Test Data

Data from both tests were collected and analyzed in two steps: (i) general analysis, and (ii) critical analysis. Microsoft Excel 2007, Spotfire Decision Site 7.3, and Observer 5.0 were used to analyse and examine these data. Due to the lack of space only a few data sets are presented here.

### 5.6.1 Data Collection

From interface sign intuitive test, data of (i) users' interpretations of all 104 interfaces signs (see table 3) as well as (ii) users' interpretations of each task-related signs were collected. From think-aloud usability test, data of (i) task completion time (TCT), min and max time for each task as well as all tasks (see table 4), (ii) number of times tried / failed to complete each task, (iii) no. of input errors, system errors as well as number of times despaired, smile, angry, asking help for each task, (iv) no. of interactions and interaction variation (difference between the interactions actually needed and user's actually performed to complete a specific task *i.e.*, *user interact - required interaction*) for each task, (v) time to stay at despaired, smiley, or angry state, (vi) time to stay at confused & wrong navigation (C&WN) state, (vii) subjective rate in the scale of 1-5 based on how easy or difficultly felt to perform each task as well as overall reaction to the studied application, and (viii) examples of verbal comments related to interface sign interpretation were collected.

Table 3. Categorizations of participants' interpretations to all interface signs

Participants	Accurate	Inaccurate				Accuracy (%)	Inaccuracy (%)
		Moderate	Conflict	Erroneous	Incapable		
P1	67	18	3	12	4	64.42	35.58
P2	65	13	8	16	2	62.50	37.50
P3	79	16	0	8	1	75.96	24.04
P4	73	14	5	4	8	70.19	29.81
P5	71	14	3	11	5	68.27	31.73
P6	77	11	2	9	5	74.04	25.96
P7	76	9	4	13	2	73.08	26.92

Table 4. TCT (mm:ss) where, min & max time cell coloured as light turquoise & rose respectively

Participants	T1	T2	T3	T4	T5	T6	T7	T8	Total
P1	01:28	05:44	08:26	00:44	05:40	02:19	01:10	00:07	25:38
P2	01:00	04:02	18:53	00:55	04:55	02:02	02:59	00:03	34:49
P3	00:38	02:57	06:47	01:44	15:18	00:29	00:36	00:03	28:32
P4	01:15	03:08	02:54	00:59	03:16	05:35	00:44	00:05	17:56
P5	01:24	06:23	07:43	00:54	01:19	06:25	01:01	00:03	25:12
P6	01:03	07:00	02:21	01:35	11:40	05:08	03:22	00:03	32:12
P7	01:12	03:11	14:07	03:04	21:27	04:39	00:14	00:03	47:57

### 5.6.2 General Analysis

Taking into account the percentage of users' interpretations accuracy and inaccuracy of (i) all interface signs as well as (ii) each task-related signs, general analysis was carried out. In this paper, the terms task-required signs refer to the signs actually needed to complete a task in a specific way and task-related signs refer to the all signs related all the distinct ways of completing a task. This analysis mainly observes the rational relation between the users' interpretations accuracy and the task completion performance (e.g., task completion time, confused & wrong navigation state, task failure, interaction variation, input error, asking help, facial expression, subjective ratings, etc.). The general analysis showed that in most of the cases users showed comparatively high task completion performance when their interpretations accuracy of interface signs was comparatively higher than others. Due to the lack of space only two examples are presented below:

*Task completion time:* Four participants *P1, P2, P4, P5* (57.14% of total participants) completed tasks in comparatively shorter time than the participants who had comparatively low interpretations accuracy of all interface signs. These participants' interpretations accuracy relation was  $P2 < P1 < P5 < P4$ , and TCT relation was  $P4 < P5 < P1 < P2$ . Again, those who have comparatively high interpretations inaccuracy of task-related signs were took comparatively more time {e.g., T2 (sign interpretation inaccuracy 45.83%, TCT 5:44) by P1, T3 (39.58%, 18:53) by P2, T4 (47.06%, 3:04) by P7, T6 (66.67%, 6:25) by P5, etc.} and vice versa {e.g., T2 (31.25%, 2:57) by P1, T3 (27.08%, 2:21) by P6, T5 (30.77%, 1:19) by P5, etc.}.

*Task failure:* Total number of task failed at the first attempt was 9. Among these, 5 tasks (55.56%) were failed by the participants where participants' interpretations of these task-related signs showed comparatively higher to inaccuracy. Tests data showed that the tasks *T3, T3, T7, T6, T6* by participants *P1, P2, P2, P4, P5*

respectively were not completed at the first attempt due to the high interpretation inaccuracy, i.e., 45.83%, 39.58%, 47.62%, 58.33%, 66.67% respectively.

### 5.6.3 Critical Analysis

Alternate cases to all categories of general analysis were also observed and these cases are named as critical cases. Few examples of critical cases are presents in below:

- Interpretation accuracy of all interface signs by P4 and P7 were 70.19%, 73.08% respectively but task completion time, interaction variation, confused & wrong navigation, despaired, angry, task failure, asking help, and subjective ratings were comparatively worse for P7 than for P4.

- Interpretation accuracy of task T5 related signs was higher for P7 than the P6 but P7 showed comparatively worse performance than P6 on the way to task completion time, interaction variation, confused & wrong navigation, despaired, asking help as well as subjective ratings.

- Interface signs for tasks T2 and T3 showed the same understandability for each participant (e.g., P1's accuracy of interpreting task-related signs for tasks T2 and T3 were 54.17%) but task completion performance differed between these tasks by each participant.

Further examination of the tests video led to analysis and discussion of these critical issues by using set operations (Set-computer science, 2010), and dependency graph (Dependency Graph, 2010). Discussions of critical cases are presented here through three observations:

*Observation I:* It was observed that the total number of interface signs of studied application was less than the number of each task-related interface signs. Again, except T1 and T8 all other tasks had more than one distinct way (e.g., to complete a task T, it is possible by interacting with the signs {S1, S2, S3, S4} and also possible by interacting with signs {S3, S4, S5}); these two sets of signs are referred as the two *distinct ways* of completing the task T) of task completion. Because of this, the number of task-required signs and number of task-related signs was not equal. Therefore, it happened that participants having comparatively low interpretation accuracy (of all interface signs as well as task-related signs) but understood the task-required signs and proceeded with those signs facilitated him to perform that task properly (e.g., low task completion time, complete the task at first attempt, low interaction variation etc.). Again, participants having comparatively high interpretation accuracy (of all interface signs as well as task-related signs) but who failed to understand the task-required signs showed worse performances for completing this task. This observation is analyzed and discussed here with an example using set operations to give a more clear idea-

*Assumptions, interface signs of whole system:*  $S = \{S1, S2, S3, S4, S5, S6, S7, S8, S9, S10, S11, S12, S13, S14\}$

*User interprets interface signs accurately for whole system:*  $A = \{S1, S5, S7, S9, S10, S11, S12, S13, S14\}$

*Task: T;* no. of distinct ways to perform task T: 4; total no. of signs for whole system:  $N = |S| = 14$

*Required signs to each ways (task-required signs):*  $W1: \{S1, S2, S4\}; W2: \{S3, S7, S8\}; W3: \{S1, S4, S5\}; W4: \{S1, S7\}$

*Thus, no. of required signs for each of these ways:*  $|W1| = 3; |W2| = 3; |W3| = 3; |W4| = 2$

*Interpretation accuracy for whole system:*  $\{(|A|/|S|) \times 100\} = 64.29\%$

*User interprets inaccurately for whole system:*  $IA = (S - A) = \{S2, S3, S4, S6, S8\} = 35.71\%$

*Related signs for task T (task-related signs):*  $RS : (W1 \cup W2 \cup W3 \cup W4) = \{S1, S2, S3, S4, S5, S7, S8\}$

*No. of related interface signs for task T:*  $|RS| = 7$

*For these, user understands accurately for task T:*  $B = RS \cap A = \{S1, S5, S7\} = 42.86\%$

*User understands inaccurately for task T:*  $C = RS - A = \{S2, S3, S4, S8\} = 57.14\%$

*Finally, these illustrate:*  $|RS| \leq N$  and  $|W1| \leq |RS| \wedge |W2| \leq |RS| \wedge |W3| \leq |RS| \wedge |W4| \leq |RS|$

$$RS \subseteq S \wedge (A \cup IA) = S \quad \text{and} \quad B \subseteq A \wedge C \not\subseteq A \wedge B \not\subseteq IA \wedge C \subseteq IA$$

$$B \subseteq RS \wedge C \subseteq RS \wedge (B \cup C) = RS \quad \text{and} \quad W1 \not\subseteq B \wedge W2 \not\subseteq B \wedge W3 \not\subseteq B \wedge W4 \subseteq B$$

That is why, though the interpretation inaccuracy (of all interface signs as well as task-related signs) were comparatively high but if user proceed with the specific way  $W4$  to perform this task T, then he/she could complete the task with high task completion performance. If user chose any other ways ( $W1$ ,  $W2$ ,  $W3$  or any other arbitrary ways) then task completion performance could be worse.

*Observation II:* The sequential as well as dependable relations were available within the set of task-required signs. For example, the dependency relation within three signs are presented as  $\{S1 \leftarrow S2 \leftarrow S3\}$ , that

means, S2 depended on S1 and S3 depended on S2 (or, to obtain S2 user needs to interact firstly with S1 and to obtain S3 user needs to interact firstly with S1 then S2). Again, the sequential relation within three signs are presented as  $\{S1 \Rightarrow S2 \Rightarrow S3\}$ , that means, S1 is sequentially related to S2 and S2 is sequentially related to S3 (or, to complete the task properly it is needed to interact firstly with S1 then S2 and finally with S3). That is why, it happened that participants having comparatively high accuracy (of all interface signs as well as task-related signs) but incapable or erroneous interpretation of one or few sign(s) within the set of task-required signs showed comparatively worse performance of task completion and subjective ratings than those who had comparatively low inaccuracy (of all interface signs as well as task-related signs). For example, to perform T3 by P2, T5 by P7 participants were not able to interpret few signs properly within the set of task-required signs; whereas these few signs were strongly related (sequentially and dependability) to other signs of that set and thus showed worse performance for these tasks. This observation is analyzed and discussed here with an example case of performing task T3 by P2 using dependency graph in a more structured way to depict the idea in a more clear way. In dependency graph (see figure 2), all circles represent related signs for task T3. The arrow sign from S2 to S1 means S2 is dependable on S1. Circles with labeling represent the required signs for the specific way of task completion, chose by P2. Set of required signs for a specific way chose by P2 for T3:  $\{S1, S2, S3, S4, S5, S6, S7, S8, S9, S10, S11, S12\}$

Dependency among these signs:  $\{S1 \leftarrow S2 \leftarrow S3 \leftarrow S4 \leftarrow \{S5, S6, S7 \leftarrow S8, S9, S10, S11\} \leftarrow S12\}$

Sequentiality among these signs:  $\{S1 \Rightarrow S2 \Rightarrow S3 \Rightarrow S4 \Rightarrow \{S5, S6, S7 \Rightarrow S8, S9, S10, S11\} \Rightarrow S12\}$

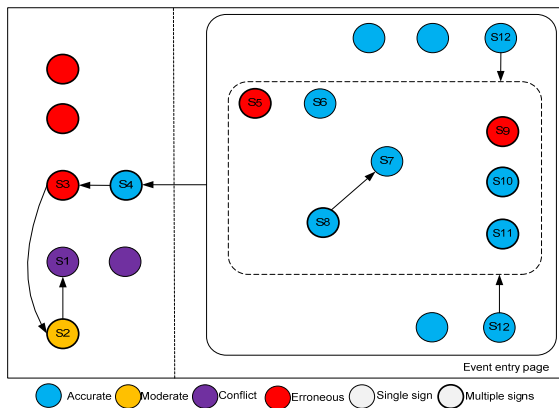


Figure 2. Sign interaction for T3 by P2

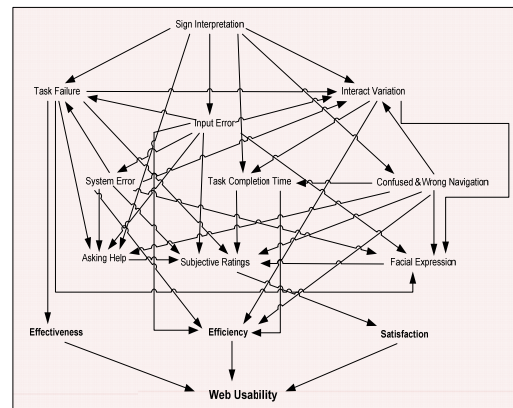


Figure 3. User's interpretations affect the web usability

That's why, to reach at event entry page dependable signs (S1, S2, S3) were not properly understandable to him thus he was tried a lot to reach at event entry page. Since, P2 needed to interact with S1 (conflicting) to reach at S2, then needed to interact with S2 (moderated) to reach at S3 and after that needed to interact with S3 (erroneous) to reach at S4 and interaction with this S4 (accurate) led him to reach at event entry page and interact with  $\{S5, S6, S7, S8, S9, S10, S11, S12\}$ . At event entry page, erroneous interpretations of S5, and S9 also affected to the task completion performance since, these signs had a dependable and sequential relation with S12. Therefore, P2 completed the task T3 with worst performance (e.g., high TCT, C&WN, task failure, high interaction variation etc.) and gave comparatively low subjective rate for this task.

*Observation III:* From this study, it was also observed that sign interpretation directly as well as indirectly affected usability metrics of effectiveness (e.g., % of goal achieved), efficiency (e.g., time to complete a task, error rate, amount of effort) and satisfaction (e.g., subjective rating scale) thus eventually affected web usability. For the lack of space only two example cases are discussed here, i.e., (i) P1 was unable to properly interpret a sign of input date (in figure 1, sign interpreted by P1) to perform task T3 and this inaccurate interpretation made an input error and this input error generated a system error and failed to perform this task at the first attempt. Then, this failure and errors showed the way of asking help twice, spend C&WN state for 32 second thus make navigation errors and these eventually directed him to increase interaction variation (51.85%). Then, these all affected to increase the TCT (8:26, whereas min time was 2:57) comparatively and finally these also affected his subjective rating (rate to 3). (ii) To perform task T5 by P7 erroneous and incapable interpretation of few signs (in figure 1, two signs interpreted by P7) led him to task failure for first

two attempts as well as spend C&WN state for 6 min 38 sec (30.91% of TCT). Then these ultimately directed him to obtain high interaction variation (1520%). After that these all affected to increase the TCT (21:27, whereas min time was 1:19) comparatively and also to his facial expressions (despaired and angry for the time of 4:09 and 0:07 respectively). Finally these affected his subjective rating (rate to 1) too. That is, an erroneous interpretation of interface signs affected directly and indirectly usability metrics (e.g., task failure affects to effectiveness; input error, TCT, interaction variation affects to efficiency; asking help, C&WN, subjective ratings to satisfaction) thus eventually affected to overall web usability. Mostly happened cases observed from this study are depicted in figure 3. Here, two nodes linked with one side arrow means arrow sided node affected by other side node. For example, inaccurate sign interpretation affected to occur input error, and this input error eventually affected to increase the interaction variation and this eventually affect to increase TCT, and so on.

## 6. RESEARCH OUTCOMES

The aim of the study was to observe usability problems in general and these problems were examined with respect to users' interpretation of interface signs in order to show the importance of considering semiotics acuity in the design and evaluation of web interfaces to boost web usability. Main outcomes of this study are presented here briefly. Users able to interpret interface signs properly complete a task in comparatively shorter time, with lowest interaction variation, and spend a shorter time on confusing and wrong navigation than users who do not understand interface signs properly. The possibility of task completion failure rate decreases with proper understanding of interface signs. Erroneous or incapable interpretation of interface signs may lead to task failure. Again, lack of proper interpretation of interface signs related to value input may lead users to make input errors. Interface sign interpretation does not affect the system error directly, but indirectly. It was also observed that ease and ability to interpret the interface signs affects users' facial expressions. Moreover, this study also showed that subjective ratings to overall satisfaction could be comparatively higher to the users who were comparatively more able to interpret interface signs properly.

This study also showed that users' task completion performances might be worse even though they had comparatively high interpretation accuracy of all interface signs as well as task-related signs. Because of, (i) number of task-related signs was greater than task-required signs ; (ii) a or few sign(s) within the set of task-required signs was (were) not properly understandable; (iii) sequential and dependable relation are also available within the set of task-required signs; and (iv) one usability problem eventually affect another problem.

A web page generally includes: content, navigation, graphics / layout, information, and interface signs. This is why, this study's objective was to depict how interface signs could affect overall usability while considering others page elements were correct from a usability perspective (see figure 3). This study showed that web interface sign presentation (design) and its interpretations affect most of the problems found through usability test. That is, sign interpretation directly as well as indirectly affects usability metrics (effectiveness, efficiency and satisfaction) thus eventually affecting web usability.

## 7. CONCLUSION AND FUTURE WORK

This paper shows the significance of semiotics perspective to web usability through an empirical case study on a web application. This study showed that semiotics consideration to interface design and evaluation were mostly important since interpretation accuracy of interface signs affect usability metrics i.e., effectiveness, efficiency, and satisfaction thus eventually facilitating optimization of the web usability. There were a few limitations to this study. Firstly, the case study was conducted only on a web application; secondly, the number of participants was rather small; and thirdly, this study did not focus on other things (e.g., content, navigation, graphics, etc.) of the web interface therefore this research did not claim that these are correctly organized in the Ovi calendar interface. The author hopes to consider these issues in future tests. Again, though many researchers have been conducted on web interface, especially on its content, navigational style, graphics/layout, and information, surprisingly web signs were always neglected. Moreover, the answer of a basic research question (*How do interface sign affect web usability?*) accomplish from this research raises



another important fundamental question “*What does the designer need to be aware of when re/design meaningful, understandable web interface signs?*” In future work, the author will seek to provide answers for this question. Therefore, this research also acted as an initial step to start my journey to work on a concrete project “*semiotics perception on web interfaces: evaluation and optimization of web usability and end user experience*”. Future research on this project will continue by focusing on interface signs, web usability, HCI, UX, and semiotics theories.

## ACKNOWLEDGEMENT

The author would like to thank all the participants of the study. Åbo Akademi Foundation (ÅAF), Finland provides the grant that has made this research possible. For this, ÅAF is gratefully acknowledged. To Franck Tétard, thank you for your feedback and suggestions for strengthening this paper.

## REFERENCES

- Al-Wabil, A. and Al-Khalifa, H., 2009. A framework for integrating usability evaluations methods: The Mawhiba web portal case study. In *IEEE International Conference on the Current Trends in Information Technology (CTIT)*, pp: 1-6, Dubai, UAE.
- Andersen, P., 1992. Computer Semiotics. In *Scandinavian Journal of Information systems*, Vol. - 4, pp. 3 -30.
- Barnum, C. M., and Dragga, S. , 2001. *Usability Testing and Research*. In Longman Publishers, New York .
- Dependency graph, 2010 . [http://en.wikipedia.org/wiki/Dependency\\_graph](http://en.wikipedia.org/wiki/Dependency_graph) (accessed on Dec, 2010).
- Diaper, D. and Stanton, N., 2003. *The handbook of task analysis for human-computer interaction*. In Mahwah, NJ: Lawrence Erlbaum.
- Frege, G., 1879. *Begriffsschrift, English translation*. In J. van Heijenoort, ed., pp. 1-82, Harvard University Press, Cambridge, MA.
- Hertzum, M., and Jacobsen, N. E., 2003. The Evaluator Effect: A Chilling Fact about Usability Evaluation Methods. In *International Journal of Human-Computer Interaction* (15:1), pp 183-204.
- Islam, M.N., 2011. "How Does User Interpretation of Web Interface Signs Affect Web Usability: An Empirical Case Study of a Web application "; *Turku Centre for Computer Science (TUCS) Technical Report*, TR No. 1000, Finland.
- Lewis, C., 1982. Using the thinking-aloud method in cognitive interface design. In *IBM Research Report RC 9265*, Yorktown Heights, NY.
- Morris, C., 1938. Foundations of the Theory of Signs. In *International Encyclopedia of Unified Science*, vol. 1, no. 2, University of Chicago Press, Chicago.
- Nielsen J. and Mack, R.L., 1994. *Usability inspection methods*. In Wiley, New York.
- Nielsen, J., 1999. *Designing Web Usability: The Practice of Simplicity*, In New Riders.
- Nielsen, J., 1993. *Usability Engineering*. In Academic Press, Inc., London, UK.
- Online resource, 2010. <http://www.webcredible.co.uk/user-friendly-resources/web-usability/user-interviews.shtml> (accessed on Nov 3, 2010).
- Peirce, C. S., 1932-58. *Collected Writings* (8 Vols.). In: Ed. Charles Hartshorne, Paul Weiss & Arthur Burks, Harvard University Press. (1931-58).
- Set - computer science, 2010. [http://en.wikipedia.org/wiki/Set\\_\(computer\\_science\)](http://en.wikipedia.org/wiki/Set_(computer_science))(accessed on Dec, 2010).
- Souze, D., 2005. Clarisse Sieckenius de Souza: *The Semiotic Engineering of Human-Computer Interaction*. In The MIT Press, ISBN: 0262042207.
- Triacca, L., 2003. Wish you were usable! How to improve the quality of a Museum Website. In *International Conference on Electronic Information, the Visual Arts and Beyond* , Florence.
- Usability Net, 2010. <http://www.usabilitynet.org/home.htm>. (accessed on Nov 8, 2010).
- Whitefield, A. et al, 1991. A framework for human factors evaluation. In *Behavior and Information Technology*, vol. 10, pp. 65-79.

# DEVELOPING USER INTERFACES FOR MONITORING SYSTEMS IN BUILDINGS

Szu-Cheng Chien, Robert Zach and Ardeshir Mahdavi

*Department of Building Physics and Building Ecology, Vienna University of Technology  
Karlsplatz 13, A-1040, Vienna, Austria*

## ABSTRACT

This paper explores the main requirements and design guidelines of user interfaces for monitoring systems in office buildings. We first discuss five information streams in the monitoring system (energy use, indoor environment, external environment, occupants' states, and environmental control systems states). We then present the results of a user survey (134 participants) and three focus group sessions (24 participants) conducted in Vienna and Taiwan. The objective of this survey was to capture the views of the potential receivers of building monitoring information regarding the relative importance of different kinds of information and the modes and means of presenting and visualizing such information. The outcome of these studies is expected to advance the state of art in connecting occupants and buildings.

## KEYWORDS

User interface, intelligent environments, monitoring systems

## 1. INTRODUCTION

### 1.1 Motivation

Energy consumption in buildings constitutes a large share of total end use energy. European Union is increasingly dependent on imported energy resources. Moreover, greenhouse gas emissions are on the rise (European Commission 2009). According to IEA (2008), buildings are responsible for approximately 40% of energy consumption and 36 % CO<sub>2</sub> emissions in the countries of the European Union. On the other hand, buildings affect health, comfort, satisfaction, and the productivity of the workforce, as over 90% of the average person's life is spent indoors. Earlier studies suggest that users' behavior and actions can influence the performance of buildings both in terms of sustainability (e.g. energy use) and indoor environmental quality (Mahdavi and Pröglhöf 2008). It is thus important to inform the users (building occupants) regarding the state of the building and its systems. Toward this end, the communication channels (and their respective interface technologies) between users on the one side and the building systems and their operators on the other side must be well-designed and effective (Chien and Mahdavi 2008).

An effective user interface platform for building automation systems has yet to emerge. Currently, most commercially available interface products are limited in functionality and effectiveness (Chien and Mahdavi 2008; Karjalainen and Lappalainen 2011). Based on advancements in ICT (Information and Communication Technologies) in recent years, new technological possibilities have emerged to better connect the occupants with environmental systems of buildings. Particularly in large and technologically sophisticated buildings, multi-faceted interactions between building occupants and the multitude of environmental monitoring devices and systems need to be appropriately aligned in order to assure effective building operation and performance. Nonetheless, relatively few systematic (long-term and high-resolution) efforts have been made to observe and analyze the means and patterns of such user-system interactions with building systems. Specifically, the necessary requirements for the design and testing of hardware and software systems for user-system interfaces have not been formulated in a rigorous and reliable manner (Chien and Mahdavi 2009). As a contribution toward this objective, this paper focuses on the emerging field of building

monitoring systems and how they could be connected with the building users. Thereby, the major requirements and design guidelines of user interfaces for building monitoring systems in office-building sector are explored. We first consider five information streams in the monitoring system (energy use, indoor and outdoor environment, occupants' states, and environmental control systems states). We then present the results of a user survey (134 participants) and three focus group sessions (24 participants) conducted in Vienna and Taiwan. The objective of this survey was to capture the views of the potential receivers of building monitoring information regarding the relative importance of different kinds of information and the modes and means of presenting and visualization such information. The outcome of these studies is expected to advance the state of art in connecting occupants and buildings. Specifically, the results can guide the process of requirement specification for user interface designs for building monitoring systems.

## 1.2 Background

### 1.2.1 Building Monitoring System

To dynamically collect and process major building-related data-streams relevant to desired user interfaces, a ubiquitous monitoring scenario must be considered. Moreover, to fully benefit from a building monitoring system, real-time sensor-data is essential (Daniels 2003). Therefore, a building communication network is usually used to automatically collect required data-streams. Such communication networks can be described with the three-layer model defined in ISO 2004. This model is appropriate to describe network communication strategies, but does not cover sensor/actor technologies. It does not deal with the challenge of getting the information from different physical domains into electronic signals and their different requirements regarding fieldbus networks. To fully cover monitoring strategies, an additional layer describing sensor/actor technologies is added. Based on the four-layer model and standardized interfaces between different layers, a mix of building communication technologies can be used. Possible technology scenarios are described in Zach and Mahdavi (2010). This adaptive approach leads to a flexible and vendor independent monitoring system, which can be optimized to address the requirements of specific buildings.

### 1.2.2 User Interface for Built Environments

An effective user interface platform for building automation systems could contribute to achieving both desirable indoor climate conditions and meeting the objectives of a sustainable building operation regime. As to the research and development concerning user interfaces in the context of intelligent built environments, there are a number of precedents. Wood and Newborough (2007) summarized the factors influencing the design of context information display for use in intelligent environments. Such factors include the place of the display, users' motivational factors, display units/methods, and timescales. Chien and Mahdavi (2009) implemented an integrated user interface system called Built Environment Communicator (BECO). It serves as a user interface model for indoor environmental controls in intelligent buildings, whereby four context categories (general information, indoor/outdoor information, and device status) and six control options/extensions (control via device, parameters, perceptual values, scenes, and schedule/micro-zoning) are considered. The ubiquitous communicator – the user interface of the intelligent house "Toyota Dream House PAPI" in Japan – is developed as a communication device that enables the occupants to communicate with people, physical objects, and places (Sakamura 2006). The HomeLab project (Philips 2008) intends to test home technology prototypes in a highly realistic way, thus speeding up technological innovations, particularly in the Ambient Intelligence domain. The MavHome (Managing an Adaptive Versatile Home) project (Cook et al. 2003), at UT Arlington, is a smart environment laboratory with state-of-the-art algorithms and protocols to provide customized, personal, safe, and energy-efficient solutions for the users. Further work on the integration of user interfaces into intelligent environments include Swiss house project in Harvard University (Huang and Waldvogel 2004), Interactive space project by SONY (Rekimoto 2009), House\_n project at the MIT Media Lab (Intille 2006), the Adaptive House at University of Colorado (Mann and Milton 2005), and many others around the world.

## 2. APPROACH

The exploration of the interface requirements for building monitoring systems involves the following three steps, namely: (i) specification of the considered evaluative categories, (ii) conducting user surveys and focus group sessions with participants, (iii) comparison of the data regarding user attitudes and preferences in terms of the specification categories.




To conduct a comparison of the data regarding user attitudes and preferences in the context of intelligent buildings, we first established a set of categories involving five data streams. These data streams include energy use, indoor and external environment, occupancy, and environmental control systems. Energy use category typically includes space heating and cooling, warm water, lighting, and equipment. Indoor environment parameters include, amongst others, room air temperature, relative humidity, air velocity, air change rate, CO<sub>2</sub> and VOC concentration (indicators of indoor air quality), and illuminance level. External environment may be characterized qualitatively in terms of weather conditions (e.g. sunny, cloudy, and rainy) or involve quantitative information on outdoor air temperature, relative humidity, wind speed and direction, as well as irradiance and illuminance. Occupant-related information includes, for example, presence and movement (number and location) of occupants and their actions. Environmental control systems (devices) information implies occupants' ability to perform queries regarding the state (position) of the building's environmental control devices. Such devices include, for example, HVAC (heating, ventilation, air-conditioning), luminaires, windows, blinds, etc.

To gain data regarding the attitudes and preferences of various stakeholders toward preferable means and modes of accessing data streams from the building's monitoring system, we conducted a user survey (questionnaire) with 134 participants (see Table 1). Thereby, the above-mentioned evaluative categories of data streams as well as user background information were considered. We included two groups, namely building experts (system developers, designers, building operators, facility managers) and building users (occupants, guests). In addition, three focus group sessions were conducted to obtain a deeper understanding of the requirement context and user experiences regarding building monitoring systems. These sessions involved 24 participants: one session was held with building experts and two sessions with building occupants. Moreover, we considered three commercially available products (see Table 2) that are designed to facilitate the communication between users and the building's monitoring system. In each session, which lasted one and half hour, we ran through five sequences, namely *i*) completing a questionnaire, *ii*) warm-up: general expression of participants' understanding of monitoring systems for buildings, *iii*) Discussion: comments on the selected interfaces, *iv*) personality profiling: game of personality mapping (McDonagh-Philip and Bruseberg 2000) with selected interface products to elicit emotional responses to products, and *v*) brainstorming: developing interface functionalities and requirements.

Table 1. User survey information (134 participants)

Gender/Marital status	Male (43%); Female (57%) / Single (68%); Married (32%)
Ages	16-20 (1%); 21-25 (22%); 26-30 (29%); 30-35 (26%); 36-40 (13%); 41-45 (3%); 46-50 (2%); 51-55 (2%); 56 and older (2%)
Residence	Austria (50%); Taiwan (50%)
Education status	High school (2%); College/University (36%); Master (53%); Doctor (9%)
Disciplinary background	Design (35%); Computer Science (22%); Management (10%); Art/Music (8%); Social Science (8%); Linguistics/Communications/Media Studies (7%); Other (10%)
User types	Occupants, guests (72%); Experts (system developers, designers, building operators, facility managers) (28%)
Smart phone ownership	54% (Austria 49%; Taiwan 59%)

Table 2. The three interface products selected for focus group discussion

Company	DIGITEXX (www.digitexx.com)	Agilewaves (www.agilewaves.com)	Oberlin (www.oberlin.edu/dormenergy/)
Illustration			

The data collected from the participants were analyzed in terms of the aforementioned evaluative categories. In addition, we considered and evaluated a number of individual statements (open-end comments) from the focus group sessions. Thereby, we focused on the following questions:

- i) What are the preferable modes of information representation and visualization regarding the aforementioned information streams from the monitoring system?
- ii) What are specific interface requirements given different types of users and their cultural background?
- iii) What are the usability implications of mobile versus stationary interface products?

### 3. RESULTS

To present the survey results in a structured manner, we use the following notation. "A" denotes all users, "B" denotes users in Austria, and "C" denotes users in Taiwan. Expert participants are specified with code "1", whereas building occupants are specified with code "2". Thus, six groups (A1, A2, B1, B2, C1, C2) are considered. As previously mentioned, we have categorized information streams in terms of energy use, indoor and outdoor environment, occupancy, and environmental control systems.

The survey results are shown in Figures 1 to 7. Figures 1 to 5 show what type of information within the aforementioned information streams was of more interest to the participants (energy use, indoor and outdoor environment, occupancy, and environmental control systems). Figures 6 and 7 provide various information regarding survey participants. Figure 6 shows what kinds of hardware would be preferred by users for accessing building-related information. Figure 7 represents users' views on the problems associated with interface systems for building-related information. Table 3 includes examples of participants' statements in the course of the group sessions.

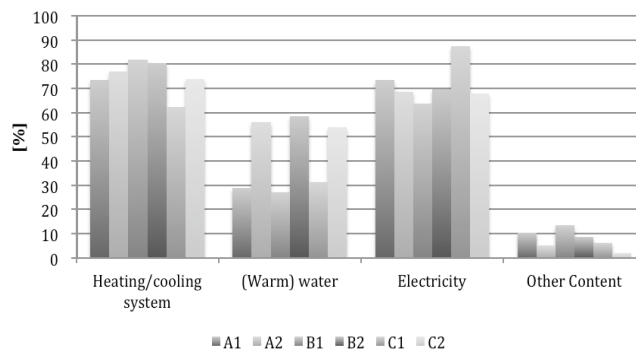


Figure 1. Participants' level of interest in energy use information of different systems

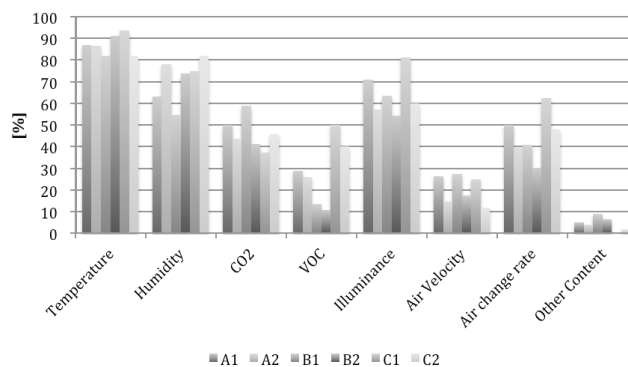


Figure 2. Participants' level of interest in indoor environment information

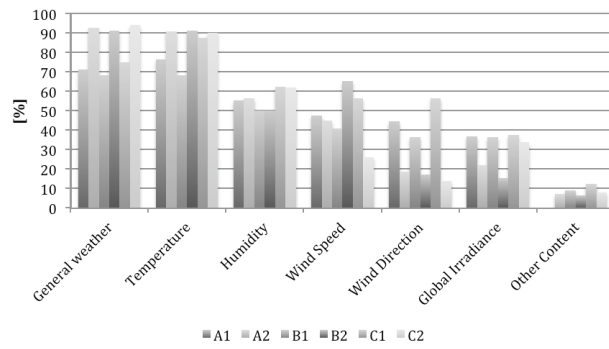


Figure 3. Participants' level of interest in outdoor environment information

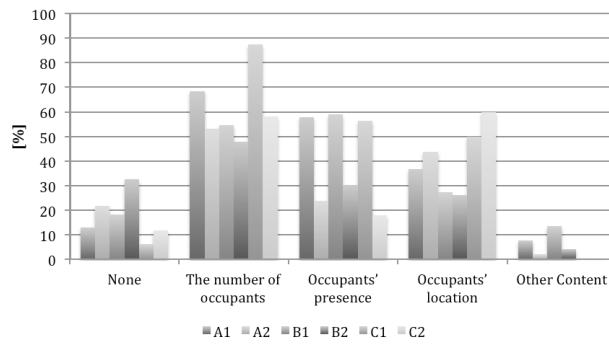


Figure 4. Participants' level of interest in occupancy information

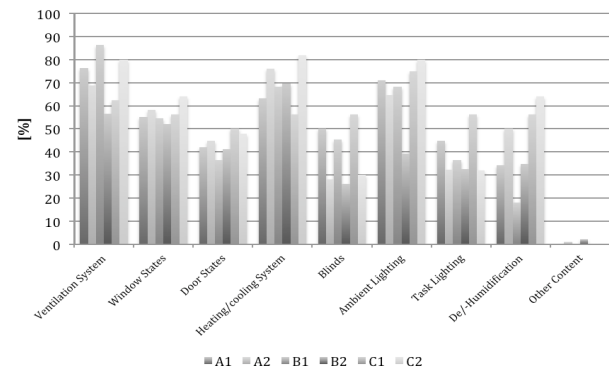


Figure 5. Participants' level of interest in building systems information

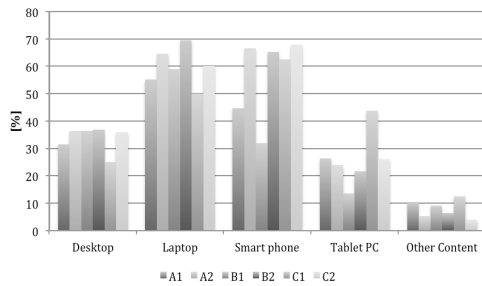


Figure 6. Participants' preference for hardware usage

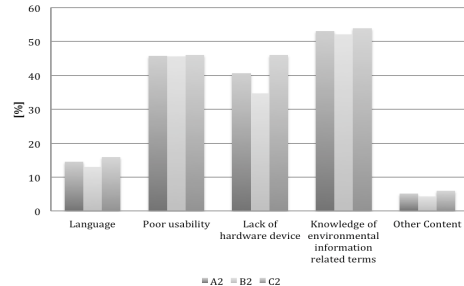


Figure 7. Participants' views on the problems associated with interface systems for building-related information

Table 3. Illustrative participants' comments in focus group session

No.	Participant	Statement
1	7	I like the "comparison" function in Product "Oberlin". Either self-comparison or comparison among buildings is an effective method and motivates me to reduce energy consumption.
2	3	We can depend on our senses to decide if we want more light, or if we want the room to be warmer or cooler. We do not need numeric information on such issues.
3	14	As a facility manager, I need to monitor building information the whole day. The ideal UI for me should take up the full screen to display more technical charts instead of (animated) icons.
4	21	I would like the UI to be clear and simple, because I only need to use it for a very short time. I found the products "Agilewaves" and "Oberlin" to be very user-friendly. Their charts, graphs, and animated icons are very easy to read. They can provide me with timely information I need.

## 4. DISCUSSION

The user survey together with focus group results support a number of initial conclusions, as discussed in the following two sections, namely functional requirements (section 4.1) and cognitive design requirements (section 4.2) of the envisioned user interface.

### 4.1 Functional Requirements

The survey results appear to suggest that "experts" would require more comprehensive technical information from a building monitoring system (see Figures 2 to 4). For example, they express more interest to obtain indoor environmental information concerning CO<sub>2</sub> and VOC concentrations as well as illuminance, air velocity, and air change rates (see Figure 2). Likewise, more interest is shown by experts regarding detailed outdoor environmental information including wind speed and direction as well as global irradiance (see Figure 3). Also, the difference in level of interest regarding occupancy numbers and presence is significant (see Figure 4). On the other hand, non-expert users express more interest in information of general character, such as indoor air temperature and humidity (see Figure 2), general outdoor weather conditions and outdoor temperature (see Figure 3). This suggests that, instead of detailed and comprehensive information levels, non-expert users should be provided with general and intuitively comprehensible information. The survey results further express a difference between users in different locations. For example, non-expert users in Taiwan show more interest in indoor environmental information (relative humidity, CO<sub>2</sub>, VOC, illuminance, and air change rates) than their Austria counterparts (see Figure 2). This observation is corroborated by the expressed interest in information concerning environmental control systems (heating, cooling, and ventilation systems, windows, ambient lighting, etc.), which is significantly higher in case of users in Taiwan (see Figure 5). This difference may be attributable to the observation, that more people in Taiwan are accustomed to climatically controlled (air-conditioned) buildings than in Austria, and thus more dependent on the proper functioning of relevant devices and their impact on indoor environmental conditions (thermal parameters, air quality, etc.).

The survey results (see Figure 1) imply a high level of interest in buildings' energy performance. A relevant question in this context may be the potential of user interfaces to not only provide energy use information, but also to motivate users toward energy efficient behavior. In the course of our focus group studies, the potential of user interfaces regarding promotion of energy-saving practices was explored (see, as an example, row 1 of Table 3). Specifically, it was noted that certain products include interesting functionalities in this regard. For example, Oberlin (see Table 2) provides comparative information on the energy performance of different buildings. It also allows for individuals to study their own energy use behavior. This suggests that inclusion of such motivating features should be considered in the design of user interfaces for building-related monitoring information. An interesting finding in our focus group sessions indicates that most users do not consider information on indoor environmental conditions essential (see, as an example, row 2 of Table 3). In fact, they consider their own perception of such conditions a reliable and sufficient source of information. However, it was very important for the users to have the possibility to

conveniently influence indoor climate in view of their preferences and needs. This would require the inclusion of control features in user interfaces for building information systems.

## 4.2 Cognitive Design Requirements

To discuss cognitive design requirements of the interface products for building monitoring systems, we consider four factors, namely mobility, usability, appropriate postures, and display format.

Mobility addresses mobile versus stationary communication devices. Specific terminals such as laptops, smart phones, and tablet PCs connected to building information model server via Internet make the concept of mobility realistic (Chien and Mahdavi 2008). As Table 1 demonstrates, almost 55% of survey participants already own smart phones. Moreover, high levels of preference are expressed for using mobile devices (see Figure 6). These observations suggest the importance of interface design strategies that properly address mobile device usage for queries pertaining to building-related information.

The survey results indicate that non-expert users identify three main problems with interface systems for building-related information: these are poor usability, hardware restrictions, and clarity of applied terminology (see Figure 7). This suggests that advanced interface products that provide non-expert users with environmental information, must also pay attention to the clarity of terms and navigational ease so that the interface of building monitoring system is easy to use and understand. In addition, the monitoring system interface should be available on a wide range of hardware devices, such that a convenient and ubiquitous access to building-related information is supported.

Posture is a way of talking about how much attention a user will devote to interacting with a product, and how the product's behaviors respond to the kind of attention a user will be devoting to it (Cooper et al. 2007). According to the qualitative discussion in our focus groups (see, as examples, rows 3 and 4 of Table 3), we concluded that the postures of the building monitoring user interface differ in view of user types. Sovereign posture application is an application involving a large set of functions and features, occupying the full screen, and using a minimal visual style. It may monopolize users' attention for long and uninterrupted periods (Cooper et al. 2007). Thus, potential users of such posture application are typically advanced user types (i.e. building experts). On the other hand, user interfaces with a transient posture must offer very short-term manipulation possibilities. This suggests that they must efficiently offer important and frequently needed functionalities, and then quickly step to background, letting the users (occupants) continue their normal activities (such as working on paper-based and screen-based tasks in offices).

It is important that the environmental information displayed in interface products is appropriate and effective for target user types. Some products (i.e. Agilewaves and Oberlin) in our study offer not only technical modes of information communication (such as charts and graphs), but also easily understandable elements such as icons and emotional pictorials and animations to present and visualize the environmental information. According to the qualitative discussion in our focus groups, non-expert users prefer such products (cp. row 4 of Table 3). On the other hand, products (such as DIGITEXX) that offer few technical representation modes (charts, graphs) appear to be preferred by expert users (cp. row 3 of Table 3).

## 5. CONCLUSION

In this paper, we explored user requirements towards desirable user interface products for building monitoring systems in technically advanced office environments. The exploration suggests an approach to context information inquiry and a set of categories for information specifications involving energy use, indoor and outdoor environments, occupancy, and environmental control systems. We have conducted preliminary user surveys together with three focus group sessions in relation to the above categories. The corresponding results highlighted a number of relative functional and cognitive design requirements, which can be further articulated toward the concrete design and implementation of a prototypical advanced interface model for monitoring systems in buildings. It is our intention to apply, for the prototyping process, the Extreme Programming software engineering method (Beck and Andres 2005), which emphasizes client-driven prototyping, usability testing, and integration cycles.



## ACKNOWLEDGEMENT

The research presented in this paper is supported in part by grants from FWF (Austrian Science Fund) project number: I 545-N22 ("Ubiquitous dynamic building performance monitoring") and the "Innovative Projekte" of the Vienna University of Technology. Also, the authors also gratefully acknowledge the support of Mr. Philip Ehrenfellner toward conducting the user surveys and focus group sessions mentioned in this paper.

## REFERENCES

- Beck, K., and Andres, D. 2005. *Extreme Programming Explained*, Addison-Wesley.
- Chien, S., and Mahdavi, A. 2008. Evaluating interface designs for user-system interaction media in buildings. *Advanced Engineering Informatics*, Vol. 22, Issue 4, pp 484-492.
- Chien, S., and Mahdavi, A., 2009. Requirement specification and prototyping for user interfaces of buildings' environmental controls. *Special Issue Building Information Modeling Applications, Challenges and Future Directions, Journal of Information Technology in Construction (ITcon)*, Vol. 14, pp 642-653.
- Cook, D. J. et al. 2003. MavHome: an agent-based smart home. *Proceedings of the First IEEE International Conference on Pervasive Computing and Communications*, pp 521-524.
- Cooper, A. et al. 2007. *About Face 3: The Essentials of Interaction Design*, Wiley, USA.
- Daniels, K. 2003. *Advanced Building Systems, A Technical Guide for Architects and Engineers*, Birkhäuser
- European Commission. 2009. *Low energy buildings in Europe: current state of play, definitions and best practice*.
- Flemming, S. et al. 2008. The Need for Human Factors in the Sustainability Domain. *Proceedings of Human Factors and Ergonomics Society 52nd Annual Meeting*.
- Huang, J. and Waldvogel, M. 2004. The swisshouse: an inhabitable interface for connecting nations. *Proceedings of the 2004 Conference on Designing interactive Systems: Processes, Practices, Methods, and Techniques*. New York: ACM Press, pp 195-204.
- International Energy Agency (IEA). 2008. *Energy efficiency requirements in building codes, energy efficiency policies for new buildings*. France. [http://www.iea.org/g8/2008/Building\\_Codes.pdf](http://www.iea.org/g8/2008/Building_Codes.pdf)
- Intille, S. S. 2006. The goal: smart people, not smart homes, *Proceedings of 4<sup>th</sup> international conference on Smart homes and health Telematics*. Amsterdam: IOS Press, pp 3-6.
- ISO 2004. ISO 16484-2, Building automation and control systems (BACS) – Part 2: Hardware
- Karjalainen, S. and Lappalainen V. 2011. Integrated control and user interfaces for a space. *Building and Environment* Vol. 46, Issue 4, pp 938-944.
- Mahdavi, A. and Pröglhöf, C. 2008. User-System interaction models in the context of building automation, *Proceedings of ECPPM conference 2008 - eWork and eBusiness in architecture, Engineering and Construction*, pp 389 - 396.
- Mann, W. C. and Milton, B. R. 2005. Home automation and smart homes to support independence, *Smart Technology for Aging, Disability and Independence: The State of the Science*. Hoboken: Wiley, pp 33-65.
- Mcdonagh-Philip, D. and Bruseberg, A. 2000. Using focus groups to support new product development. *Institution of Engineering Designers Journal* Vol. 26, pp. 4-9.
- Philips 2008. *HomeLab- Our testing ground for a better tomorrow*. Retrieved 8 2008, from <http://www.research.philips.com/technologies/misc/homelab/>
- Rekimoto, J. 2009. SenseableRays: Opto-haptic Substitution for Touch-enhanced Interactive Spaces, *Proceedings of ACM CHI 2009 conference on Human Factors in Computing Systems 2009*. New York: ACM Press, pp 2519-2528.
- Sakamura, K. 2006. Challenges in the age of ubiquitous computing: a case study of T-Engine, an open development platform for embedded systems. *Proceedings of the 28th international conference on Software engineering*. New York: ACM Press, pp 713-720.
- Stern, P. C. 1999. Information, incentives, and pro-environmental consumer behaviour, *Journal of Consumer Policy* 22, pp 461-478.
- Wicker A. 1969. Attitudes versus actions: The relationship of verbal and overt behavioral responses to attitude objects. *Journal of Social Issues* 25, pp 41-78.
- Wood, G. and Newborough, M. 2007. Energy-use information transfer for intelligent homes: enabling energy conservation with central and local displays, *Energy and Buildings*, Vol. 39, Issue 4, pp 495-503.
- Zach, R. and Mahdavi, A. 2010. Monitoring for Simulation Validation. *BauSim 2010 - Building Performance Simulation in a Changing Environment*. Vienna, Austria, pp 190-195.

# READING AS PLAYING: A NEW TUTORING MULTIMEDIA TOOL FOR CHILDREN WITH TEXT COMPREHENSION PROBLEMS

Michela Carlini<sup>1</sup>, Tania Di Mascio<sup>2</sup> and Rosella Gennari<sup>3</sup>

<sup>1</sup>Free University of Bozen-Bolzano, viale Ratisbona, 16, I-39042, Italy

<sup>2</sup>DIEI – University of L'Aquila, Via G. Gronchi – I-67100, L'Aquila, Italy

<sup>3</sup>Free University of Bozen-Bolzano, Piazza Domenicani 3, I-39100 Bolzano, Italy

## ABSTRACT

Nowadays, several children have deep text comprehension problems, despite well developed low-level reading skills. We are working on a tutoring multimedia tool aiming at implementing a series of reading intervention for such children through smart game. This paper reports on the main choices of our tool, and results of our evaluations.

## KEYWORDS

Supporting User Populations with Physical Disabilities. Graphical Interfaces. Usability.

## 1. INTRODUCTION

Developing the capabilities of children to comprehend written texts is key to their development as young adults. Text comprehension skills and strategies develop enormously from the age of 7-8 until the age of 11, when children develop as independent readers. Nowadays, more and more children in that age range demonstrate difficulties in deep text comprehension, despite well-developed low-level cognitive skills like vocabulary knowledge. Several studies experimentally demonstrate that these children with deep text comprehension problems fail to master the following reasoning skills in processing written stories, skills that are causally implicated in the development of deep text comprehension: (s1) coherent use of cohesive devices such as temporal connectives, (s2) inference-making from different or distant parts of a text, integrating them coherently, and (s3) detection of inconsistencies in texts.

For instance, see (Cain and Oakhill, 1999) for the case of poor comprehenders, circa 10% of the 8-10 olds without physical disabilities, and see (Marschark et al., 2009) for the case of deaf children. In particular, experiments show that inference-making questions centred around (s1), (s2), and (s3), together with adequate visual aids, are pedagogically effective in fostering deep comprehension of stories, e.g., see (National Reading Panel, 2000).

We are working on a tutoring multimedia tool for hearing poor comprehenders and deaf children, aged 8-10, that fail to master the above reasoning skills. Stories, adapted to the specific requirements of these children, constitute the reading material of our tool (currently, in Italian). Like Intelligent Tutoring Systems (ITSs), our tool aims at producing learning gains. However, students often dislike interacting with ITSs. According to several authors, e.g., (McNamara et al., *in press*), a potential remedy lies in games. Nowadays children are used to multimedia environments (albeit not all are equally skilled in them), and approach learning with different expectations than children of 20 years ago. This is particularly the case of our tool's users: according to the experts we interviewed, our end users like playing videogames that motivates them more than print reading. Our tool, therefore, foresees interactive games for engaging children in reading, deeply, stories. The games have questions centred around reasoning skills like (s1), (s2), and (s3) above for fostering the texts' deep comprehension. The currently games invite readers to reason on the main events of a story, and correlate them by means of temporal cohesive devices like "before", "while" or "after". In this manner, the games will help readers in constructing their mental model of the flow of the story's events.

Based on the seminal work of (Pavio, 1991), numerous studies already showed significant comprehension

gains when people can appropriately visualize while reading, as reported in (Johnson and Glenberg, 2005).

Our tool aims to be visual as for: the global interface, the story's main events, and the games. This paper reports on the main choices of our learning multimedia tool, and results of our evaluations.

## 2. THE USER INTERFACE

The Graphical User Interface (GUI) of our tool is designed following the focus+context method and developed in Adobe Flash 4. The GUI aims at implementing the tool's main functionalities: reading the stories, playing the games, and analysing the difficult words of the story. The user is invited to read a story; after reading and possibly analysing the story's difficult words, the user can choose to play the tool's interactive smart games that pose different types of questions for reasoning over the read story.

As our main goal is enjoying the stories' reading and playing experience, the GUI design emphasizes the role of the stories' illustrations whereas all the other design elements of the GUI (e.g., help) are sober and neutral. The use of metaphors is very limited and there is no use of anthropomorphic elements other than the characters contained within the illustrations to the stories. In this manner, we aim at enhancing the visual strength of the illustrations for the stories and games. Following the same principle, a textual help guide was preferred over an agent in order to avoid any possible visual clash or competition with the reading or playing activities. Moreover, children can get as quickly bored with artificial characters as adults do, and tend to ignore their instructions (Preece, 2002), whereas this does not happen as quickly with textual instructions. Sound is not used, since also profound deaf children are among the end users of our tool.

Since the cognitive processes in young children rely on a direct experience of reality, and more specifically on a complete sensory and motor perception of space and objects (Bernardinis et al., 1994) the GUI of our tool aims at offering a perceivable environment in its entirety through a reduced number of interaction modalities: the *exploring*, the *story reading*, the *word analysing* and the *game playing* modalities.

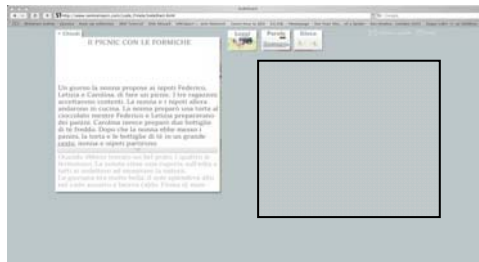


Figure 1. Page Structure: the text (the context) in the left panel is always displayed, the right panel contains the focus (illustrations, words, and videogames) that changes in relation to the modality chosen by the user.

The tool's modalities are realized in the *entrance page*, the *story page*, the *word page*, and the *game page*. The last three pages are in turn composed of two panels: the left panel and the right panel. The general structure of the page is shown in Figure 1. The left panel of the story, the word, and the game pages is the *text* panel, implementing the reading modality, as that figures shows. The right panels of those pages have different purposes and contents that depend on the page: the *illustration* for the story page, the *vocabulary* for the word page, and the *videogames* for the game page. So, the right panel of the word page implements the *word analysing* modality, and the right panel of the game page implements the *game playing* modality. All pages implement the *exploring* modality. Such choices realize the focus+context design of the GUI. Except for the entrance page, the other pages always display the context (that is the story text), and the changes of focus in relation to the modality. Modalities, pages and panels are described at length in what follows.

### 2.1 The Exploring Modality

The exploring modality, present in all pages, serves to: (1) browse among the stories; (2) browse among the story' words of the vocabulary; (3) browse among the games; (4) browse the help; (5) discover the two types of panels; (6) navigate the other three modalities: the story, the vocabulary and the game modalities.

In the case of (1), (2), and (3), the user can browse stories, words and games fading in/fading out their

respective representations: iconic stories titles in the entrance page (see Figure 2), story' words in the right panel of the word page (that is the vocabulary panel, see Subsection 2.3.1), and iconic names of games in the right panel of the game page (that is the videogame panel, see Subsection 2.4.1). Once the user chooses a story, s/he can choose another one just closing the first and returning to the entrance page; operatively, s/he clicks on the “chiudi” (close) link positioned in the top left part of the pages, e.g., see Figures 4 and 5.



Figures 2. In the left part. Entrance Page: fading in/fading out stories.

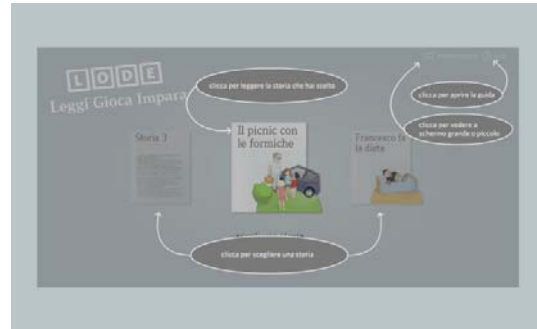


Figure 3. In the right part. Entrance Page: the Help.

The help is designed as a transparent layer that displays on top of the applications and is accessible at any time with one click on the “aiuto” (help) link positioned in the top right part of the GUI pages. The help presents balloon-like instructions written in a conversational, children-friendly style, e.g., see Figure 3. More in general, each page has its own help that, for coherency, is kept as a simple textual guide through all modalities. It is sufficient to click on any part of the GUI in order to exit the help.

The default choice of the screen is the small screen; to go in the full screen mode, the user can easily click on the “schermo grande” (big screen) link; otherwise, if the user is in the full screen mode, the GUI returns in the small screen mode by clicking on the “schermo piccolo” (small screen) link.

In order to explore the story, word and game modalities, the users have at their disposal corresponding navigation links, labelled “leggi” (read), “parole” (words), and “gioca” (play). They are displayed in a top bar as boxes, with graphical representations. The user accesses a modality by clicking on the link in the corresponding box, e.g., the user accesses the game modality by clicking on the “gioca” (play) link. This box remains of small size while the user remains in the chosen modality. See Figures 4 and 5.



Figures 4. In the left part. Word Page: the closed box of the navigation links is relative to the Word modality



Figure 5. In the right part. Game Page: the closed box of the navigation links is relative to the Game modality.

## 2.2 The Story Reading Modality

These modality presents a set of stories. Once the user chooses a story in the entrance page, the story is displayed in the story page. The left panel of the story page is the text panel: it shows the text of the story. The right panel of the page is the illustration panel: it shows the illustrations of the chosen story. In the left panel, the text is chunked into paragraphs, each corresponding to an episode of the story. Each episode has its

own illustration. This is placed on a spatial map shown in the illustration panel to the right (Figures 6 and 7). The episode in the centre of the text panel is displayed in a frame delimited by two arrows, a top and a bottom grey arrows. The frame works as a sort of lens for highlighting the episode. The user moves from the highlighted episode to the adjacent episodes with these scrolling. Correspondingly, the right panel moves through the map and zooms in on the illustration that shows the currently highlighted textual episode. The episode-stream in the text panel is synchronized to the animation in the illustration panel. Nevertheless it is the user who determines the pace of streaming accordingly to his/her needs in reading in the text panel.

### 2.2.1 The Text Panel

Textual navigation in the left text panel is constrained and over simplified primarily for avoiding any potential cognitive overload other than that due to text reading (Salmerón and García, *in press*). In particular, each story is displayed in one single page for allowing the user to gain an immediate global overview of the story's length and structure, unlike what would happen with a text displayed through several pages as in LODE (Gennari and Mich, 2007). Adjacent episodes are displayed next to the highlighted episode for simplifying the memorization of the ordering of events in the text, and hence easing the user's orientation within the text. Such a persistence of context allows the user to move back and forth between difficult episodes several times, viewing them always embedded in their surrounding context.



Figures in the left part. The Story Page: the second episode in the text and its own illustration in the illustration panel.



Figure 7 in the right part. The Story Page: the fifth episode in the text panel and its own illustration in the illustration panel.

### 2.2.2 The Illustration Panel

According to our expert-based evaluations (see Section 3.1 below), the illustration of a textual episode should not become a shortcut for the comprehension of the text: story-reconstruction through the viewing of the images has to be made impossible so that the user is compelled to read the text. Therefore, illustrations do not present any visual clue concerning the temporal flow of the stories on purpose. At the same time, the visual component of the application must be appealing and comply with the standard of printed books for children, where illustrations function as memory-reinforcement and attention-catalysts. Then, the illustration of an episode characterizes the actors and the spatial locations of the episode's main events (Figures 6 and 7).

The movements from one illustration to the other are rendered as camera movements over the spatial map through sliding, zooming in, and zooming out effects between freeze frames of single locations on the main scene. The global view over the map allows for the direct perception of the whole narration; the narrative space is thus directly perceivable as a physical space where the illustrated episodes are physically located. This creates a perceivable correspondence between the textual episode and its illustration, between the user's movement through the text and the camera movements in the map. Notice that animation is used exclusively to display camera movements. As such, the animation has the precise function of attracting the user towards the story's episodes, their actors, and their spatial locations. Such a spare use of animation effects has to be ascribed to their ambivalent potential, both in attracting attention, but also in unwanted power to distract the user from the priority action, which is reading. Please note that the illustration in the navigation link box "leggi" (read) is coherent with the episode's illustration in the right panel.

## 2.3 The Word Analysing Modality

This modality, implemented by the word page, presents the meaning in Italian and in LIS (Italian Sign Language) of a set of all the potentially difficult words of the stories for the intended end users. Once the user clicks on the navigation link “parole” (words, see Section 2.3.1), or on the highlighted word in the text panel, the word page appears. The left text panel of the word page shows is the text panel; the right panel is the vocabulary panel and shows the meaning of the clicked difficult word.

### 2.3.1 The Vocabulary Panel

This panel presents, in the low part, the textual meaning of the highlighted word, and, in the top part, the LIS video translation of the clicked word with explanations and usage examples (see Figure 4). As in the case of the illustration panel, also in this case, the contents of this panel changes coherently with the choices made in the text panel. There is a complete simultaneity in following the stream in the text panel through its difficult words and their translation in the vocabulary panel.

## 2.4 The Game Playing Modality

The game playing modality of the game page presents three types of question-games in the currently implemented version. The implemented questions address specific features of events (e.g., who bakes the cake?, see Figure 8), and causal-temporal relations between events (e.g., do the ants eat the picnic before grandma returns from her stroll?). They are based on comprehension interventions centred around inference-making in order to improve the text comprehension skills of the intended end users. See Section 1.

Once the user clicks on the navigation link “gioca” (play, see Section 3.1), the game page appears. Again, the left panel of the game page is the text panel, consistently with the other modalities. The right one is the videogame panel. The narration is kept available all the time to be read and got over again during game playing. In fact, children solve logical operations more easily on material they can perceive directly through the senses; as soon as they have to operate on an abstract level their failure rate tends to increase.

### 2.4.1 The Videogame Panel

The videogame panel, to the right, proposes three types of games concerning the read story. Each type is rendered with its own iconic representation.

The games are in a videogame format that is both familiar and attractive to its users (see Figure 8). The user plays and accumulates points according to his/her answers to the game-question. The scores are always available as feedback, so that the users can monitor their text comprehension and are also encouraged to gain more points (see Figure 9). The written story is always available in the left panel and can be scrolled with the scrolling arrows while playing. In this manner, our user can always reread a passage of the written text, operate logically on the text, and answer a question in the playing area in the right panel.



Figure 8. in the left part. The Game Page: the games questions.



Figure 9. In the right part. The Game Page: the final score of the game

### 3. EVALUATION

As recommended in the user-centred design methodology (UCDM) (Norman and Draper, 1986) the conceptual design of the tool developed through evolutionary prototyping from the beginning throughout the whole project. The prototyping phase lasted circa twelve months (June 09–August 10) and comprised low-fidelity (paper sketches, storyboards) and high fidelity versions built in adobe flash. The version presented here is the latest flash-prototype, developed in September 2010 after going through several iteration cycles and evaluations. The evaluations concerning the usability of the tool were performed using expert-based methods (the cognitive walkthrough method (Wharton et al., 1994)) and user-based methods (observational evaluation and verbal protocol (Hartson et al., 2001)). Hereby we recap the results of the main evaluations.

#### 3.1 Expert-based Evaluations

Following the cognitive walkthrough method, the designer conducted separate evaluation sessions about specific issues with experts concerning: the illustrations of the stories, and their function with two experts of children's literature; the overall design choices (e.g., typography) with an expert of multi-media communication and psychologists expert of our end-users; the usability of the tool with an expert of usability.

The expert-based evaluations with the experts of our end-users made us choose not to illustrate temporal features of the stories in the right panel of the story modality, as explained above. The subsequent evaluations with the experts of story illustrations were all positive concerning the realized illustrations. The evaluations with usability experts served to resolve predictable usability problems.

#### 3.2 User-based Evaluation

These evaluation aimed at detecting further usability problems and assessing the user satisfaction. In particular: *Assessment Goals*: Assessment of – G1: text usability (colour, font, link); ,G2: links quality; G3: the help interaction; G6: browsing among modalities; G6: games interaction; *Coherency Goal*– G5: coherency between the illustrated episode and the textual episode; *Satisfaction Goals* – Satisfaction in – G7: playing games; G8: reading stories; G9: reading word' definitions.

The sessions of this evaluation were conducted based on a classical HCI user based schema, e.g., see (DiMascio et al., 2005). The methods are direct observational evaluation methods and verbal protocols.

##### 3.2.1 User Analysis

Our experiment participants were 3 children aged 8–10 year old. *User A*: female, 10 year old; hearing; medium degree of text comprehension; low attitude to reading book; high attitude to see cartoons; low attitude to playing videogames. *User B*: male, 8 year old; hearing; low degree of text comprehension; low attitude to reading book; high attitude to watching cartoons; high attitude to playing videogames. *User C*: female, 9 year old; hearing; medium degree of text comprehension; medium attitude to reading book; high attitude to watching cartoons; medium attitude to playing videogames.

##### 3.2.2 Experiment Design

In order to better observe users, we decide to make one session per users. Each session is divided into four phases, one per modality: Phase (1), exploring modality, addressing goals G3 and G6; Phase (2), the text reading modality, addressing goals G1, G4, G5 and G8; Phase (3), the word analyzing modality, addressing G9; Phase (4), the playing game modality, addressing goals G4, G6, and G7. For each phase we defined different tasks, listed as follows. Note that  $T_{ij}$  is the task  $i$  of phase  $j$ .

*Phase 1*: T1,1: choose the “Francesco e la Dieta” story; T1,2: choose the “il picnic con le formiche” story; T1,3: ask for help; T1,4: close the help; T1,5: go to the game modality; T1,6: go to the vocabulary modality; T1,7: read another story; T1,8: quit the system. *Phase 2*: T2,1: read the story; T2,2: read the third episode. *Phase 3*: T3,1: read the definition of “furbo” (cunning) word; T3,2: see the LIS video of the “furbo” word. *Phase 4*: T4,1: play game “Salva i panini” (Save sandwiches); T4,2: play game “Calcio” (Football); T4,3: once the game ends, choose another game.

### 3.2.3 User Teaching

Before performing the experiment, the evaluator met the children's mothers. The evaluator discussed the organization of the experiment (e.g., meeting time, sequence and nature of tasks), and their respective roles in the experiment; e.g., mothers were asked not to support their child in any of the phases of the experiment.

### 3.2.4 Experiment Execution

We conducted the experimental sessions during the period of August 2010, in different dates (Users A,B,C on the 8th 10th, 18th of August, respectively). The location was the user's houses. The evaluator chose to start with the observational evaluation methods. At the end of each phase the evaluator asked the children questions about their mistakes, and indirect questions concerning their satisfaction, e.g., "would you like to play again this game or not?". In the following, we summarize the order of tasks per phase and per session user and the time spent for each session:

- User A – Phases 1, 2, 3, 4 – Tasks T1,1; T1,3; T1,4; T1,5; T1,7; T2,1; T2,2; T3,1; T3,2; T4,1 – Time: 1h.
- User B – Phases 1, 2, 3, 4 – Tasks T1,2; T1,3; T1,4; T1,6; T1,5; T1,8; T2,2; T2,1; T3,2; T4,2 – Time: 30m.
- User C – Phases 1, 2, 4, 3 – Tasks T1,2; T1,3; T1,4; T1,7; T1,8; T2,1; T2,2; T3,2; T3,1; T4,2 – Time: 1h.

### 3.2.5 Results Analysis

For space limitations, this section only gives the most significant results for the design of the GUI. These results are presented phase per phase.

*Phase 1* – 1a) Tasks T1,1,2,8 – When the users interact with the entrance page, all of them easily choose their story and two of them positively comment on the illustrations. All users easily quit from the system. 1b) Tasks T1,3,4 – The Help is easily opened by all, but two of them (Users A and C) ask the evaluators on how to exit. 1c) Tasks T1,5,6 – Once the users are in the story reading modality, all of them can easily change to another modality. 1d) Task 1,7 – One of the users (User b) clicked on "chiudi" (close) to change the story. Others click on "leggi" (read), asking the evaluator why the system does not work as they would expect it.

*Phase 2* – 2a) Tasks T2,1,2 – All users complete the tasks, but, when a user can preliminary read the third episode (T2,1) and then all the story (T1,1), he/she asks the evaluator about the location of the third episode. When the order of the tasks is T2,1 and then T2,2, all the users correctly use the scrollable arrows.

*Phase 3* – 3a) Tasks T3,1,2 – All users easily complete the tasks, but none of them understand that an underlined word is to be explained in the word analysing modality.

*Phase 4* – 4a) Tasks T4,1,2 – All users easily choose the games and just one of them (User A) asks the evaluator which games to play (Which, When or Where), the others independently choose any of the games. 4b) Task 4,3 – Not all users complete this task. All of them ask about the button to click for playing again.

### 3.2.6 Short Discussion

Specific usability issues resulted from the described user based evaluation. For instance, the GUI of our tool needs a more prominent and explicit link to exit from the help. The typography of the text needs improvement, e.g., larger fonts; in particular we will evaluate other types of scrolling arrows, more evident than the evaluated ones. Moreover we should also more strongly highlight the word of the vocabulary in the left text panel and the game exit needs to be more evident.

## 4. CONCLUSIONS

This paper presented and motivated the choices of our tutoring multimedia tool for 8-10 olds with deep text comprehension problems. The tool has illustrated stories and games for reasoning about specific features of its stories. Its games render typical reading interventions for our users in a playful format using the illustrations of the stories. More precisely, the games are in videogame format, which is appealing for our end users. Expert based evaluations were crucial for assessing the best type and role of illustrations for stories, and for choosing the videogame format for the tool's games. The user based evaluations, serving to detect and resolve usability problems, assess the satisfaction of our end users.



## ACKNOWLEDGEMENT

The research of Di Mascio and Gennari was supported by the European Community's Seventh Framework Programme FP7/2007-2013 under the TERENCE grant agreement n. 257410. Gennari work was also supported by a grant of the CARITRO bank of Trento for the LODE project.

## REFERENCES

- Bernardinis, M., Costa, R., and Galliani, L. (1994). *Immagine continua: Mappe cognitive e percorsi curricolari dai 3 ai 14 anni: Spazio e tempo nell'esperienza audiovisiva*. Padova: Cleup.
- Cain, K. and Oakhill, J. V. (1999). *Inference Making and its Relation to Comprehension Failure. Reading and Writing*, Vol. 11, pp 489-503.
- Di Mascio, T., Catarci, T., Santucci, G., Dongilli, P., Franconi, E., and Tessaris, S. (2005). Usability Evaluation in the SEWASIE project. *In Proc. of HCI2005*. Las Vegas, USA. Springer.
- Gennari, R., and Mich, O. (2007). Constraint-based Temporal Reasoning for E-learning with LODE. *In Proc. of CP 2007*, pp. 90-104. LNCS, Springer.
- Hartson, H., Andre, T., and Williges, R. (2001). Criteria for Evaluating Usability Evaluation Methods. *International Journal of Human Computer Interaction*, Vol. 13 No. 4, pp. 373-410.
- Marschark, M., Sapere, P., Convertino, C., Mayer, C., Wauters, L. and Sarchet, T. (2009). Are Deaf Students' Reading Challenges Really About Reading? *American Annals of the Deaf*, Vol. 154, No 4, pp 357-367.
- McNamara, D.S., Jackson, G.T., and Graesser, A.C. (in press). Intelligent Tutoring and Games (ITaG). *Y.K. Baek (Ed.), Gaming for Classroom-based Learning: Digital Role-Playing as a Motivator of Study*. IGI Global.
- National Reading Panel (2000). Teaching Children to Read: An evidence-based Assessment of the Scientific Research Literature on Reading and its Implications for Reading Instruction. Washington, D.C. *National Institute of Child Health and Human Development*.
- Norman, D.A. and Draper, S.W. (1986). *User Centered System Design: New Perspectives on Human-Computer Interaction*. Lawrence Erlbaum Associate. Hillsdale.
- Paivio, A. (1991). Dual-coding Theory: Retrospect and Current Status. *Canadian Journal of Psychology*, Vol. 45, pp 255-287.
- Preece, J. (2002). *Interaction Design : Beyond Human Computer Interaction*. New York, NY: Wiley.
- Salmerón, L., and García, V. (in press). *Reading skills and children's navigation strategies in hypertext. Computers in Human Behaviour*.
- Wharton, C., Rieman, J., Lewis, C., and Polson, P. (1994). The Cognitive Walkthrough Method: a Practitioner's Guide. In J. Nielsen, and R. Mack (Eds), *Usability Inspection Methods* (pp. 105-140), John Wiley and Sons, Inc..

# USING GOOGLE ANALYTICS TO EVALUATE IMPROVEMENTS IN THE USABILITY AND PERFORMANCE OF E-COMMERCE WEBSITES

Layla Hasan

*Department of Management Information Systems, Zarqa University  
Zarqa, Jordan*

## ABSTRACT

Despite the importance of good usability in e-commerce websites, few studies were found in the literature that evaluated the usability of such sites. Those that were found focused on the identification of usability problems and provided suggestions on how to improve the usability of the sites under investigation. Little research, however, investigated the effect of improving the usability of an e-commerce website in terms of its web usage and performance. This research used a matrix of advanced web metrics, calculated using Google Analytics software, to evaluate the improvements in usability of an e-commerce website after being redesigned based on an initial evaluation of its usability. The research showed that the matrix of advanced web metrics was a useful tool to illustrate the improvements of a site's usability and performance. The research also illustrated the significant influence of improving the usability of an e-commerce website on its web usage and performance. However, because the metrics provided potential indications of usability issues, a further research is required. The next step will be to employ user testing method to confirm the improvement indications provided by the metrics.

## KEYWORDS

Usability testing, Google Analytics, web analytics, user testing, e-commerce websites.

## 1. INTRODUCTION

Usability is one of the most important attributes of any user interface and measures how easy the interface is to use [11]. Nielsen and Norman [12] stressed the importance of making e-commerce sites usable. They do not regard good usability as a luxury but as an essential characteristic if a site is to survive.

Research has offered some advantages that can be gained if the usability of e-commerce websites is considered or improved. Nielsen and Norman [12] indicated that addressing the usability of sites could increase the percentage of visitors who purchased from a site and who could then turn into frequent and loyal customers.

A variety of usability evaluation methods has been developed to identify usability problems. The most well-known usability evaluation methods can be classified into three categories in terms of how the usability problems are identified: by users, evaluators or tools:

- **Evaluator-based usability evaluation methods:** This category includes usability methods that involve evaluators in the process of identifying usability problems. Examples of common usability methods related to this category are: heuristic evaluation, pluralistic walkthrough and consistency inspections.

- **User-based usability evaluation methods:** This category includes a set of methods that involves users. These methods aim to record users' performance while interacting with an interface and/or users' preferences or satisfaction with the interface being tested. The most common method in this category relates to user testing, which provides direct information regarding how real users use the interface and illustrates exactly what problems users encounter in their interaction [13].

- **Tool-based usability evaluation methods:** This category includes software tools which automatically evaluate the usability of a website. An example of these tools includes web analytics tools. Web analytics is

an approach that involves collecting, measuring, monitoring, analysing and reporting web usage data to understand visitors' experiences [10, 17].

Research concerned with the evaluation of e-commerce websites focused on the identification of usability problems and provided suggestions to improve the usability of the websites. However, little research has been conducted to investigate the usability, performance, and web usage of the websites after making the suggested changes. The research described here aims to address this gap and presents the results of investigating the usability of an e-commerce website after implementing changes and improving its usability. This paper is organised as follows: Section 2 describes web analytic tools and web metrics, Section 3 reviews related work, Section 4 presents the aims and objectives of this research, Section 5 describes the methods used, Section 6 presents the main results, and finally, Section 7 presents some conclusions.

## 2. WEB ANALYTICS AND WEB METRICS

Web analytics tools can be categorised in terms of their web traffic data source. The two most common methods used by web analytics tools to collect web traffic data are: the server-based log file and the client-based page-tagging (JavaScript tagging) approaches. The server-based log file method was the first data source used by web analytics tools and it involves the use of a server's log file to collect data [9].

Kaushik [8] indicated that while the log file technique was used widely as a data source for web analytics, the disadvantages of using this approach (i.e. *page caching*, *inaccuracy in identifying unique visits*) were noticed by both web analytics vendors and customers. These challenges led to the emergence of page-tagging techniques as a new method of collecting data from websites. The page-tagging (JavaScript) technique involves collecting information by page view and not by hits, which is the method used by the log-file [14]. It involves adding a few lines of script (JavaScript code) to the pages of a website to gather statistics from them. The data are collected when the pages load in the visitor's browser as the page tags (JavaScript code) are executed. The JavaScript tagging approach has several advantages including: accuracy, which is due to several reasons. For example, the data are collected directly from the users and not from the web server; also, most page tags that determine the uniqueness of a visitor are based on cookies (this method is not influenced by the cache technique because it collects information from every page as the code is executed every time the page is viewed, regardless of where the page was served) [8, 14].

An example of a web analytics tool that uses the page-tagging approach and which has had a major effect on the web analytics' industry is Google Analytics (GA) [8]. In 2005 Google purchased a web analytics firm called Urchin software and subsequently released Google Analytics to the public in August 2006 as a free analytics tool.

Web metrics are employed to give meaning to web traffic data collected by web analytics tools. Web metrics can be placed into two categories: basic and advanced. Basic metrics are raw data which are usually expressed in raw numbers (i.e. visits). Advanced metrics are metrics which are expressed in rates, ratios, percentages or averages instead of raw numbers, and are designed to guide actions to optimise online business. Inan [7] and Phippen *et al.* [15] criticised the use of basic metrics to measure the traffic of websites. Instead, they suggest using advanced metrics.

## 3. EVALUATING THE USABILITY OF E-COMMERCE WEBSITES

Despite the importance of good usability in e-commerce websites, few studies were found in the literature that evaluated the usability of such sites. Those that were found employed usability methods that involved users, evaluators or web analytics tools in the process of identifying usability problems. Tilson *et al.*'s [16] study is one that involved users in evaluating the usability of e-commerce websites. The researchers asked sixteen users to complete tasks on four e-commerce websites and report what they liked and disliked. Major design problems encountered by users while interacting with the sites were identified and, based on them, the researchers provided suggestions for improving the usability of e-commerce sites. The study conducted by Freeman and Hyland [4] also involved users in evaluating and comparing the usability of three supermarket sites that sold multiple products. The results proved the success of the user testing method in identifying various usability problems on the three sites.

Other studies have involved evaluators using the heuristic method to evaluate the usability of e-commerce websites [2]. Chen and Macredie [2] employed this method to investigate the usability of four electronic supermarkets. The results demonstrated the usefulness of the heuristic evaluation method regarding its ability to identify a large number of usability problems (weaknesses) and a large number of good design features (strengths) of the sites. Barnard and Wesson [1] employed both the user testing and heuristic evaluation methods together to identify usability problems for e-commerce sites in South Africa from the perspective of both experts and users. Several usability problems were identified on the selected South African e-commerce sites; this proved the success of these methods in identifying a comprehensive set of usability problems.

Web analytics tools, specifically Google Analytics, have also been used to evaluate the usability of e-commerce websites. Hasan *et al.* [5] conducted research that involved three e-commerce case studies where a comparison was made between the usability findings indicated by Google Analytics software and results from a heuristic evaluation of the sites conducted by web experts. The research identified and suggested specific advanced web metrics, calculated using Google Analytics software, which can be used to indicate general usability problem areas and specific pages in an e-commerce site that have usability problems. Hasan *et al.* [5] concluded that the suggested web metrics cannot provide in-depth details about specific problems that might be present on a page, and therefore they recommended using other usability techniques, such as heuristic evaluation, in order to obtain a more thorough appreciation of the issues.

The literature outlined above indicates that there is a lack of research that investigates the usability of e-commerce websites after correcting the usability problems identified by a usability evaluation method to check whether their usability, web usage and performance have been improved. It is worth mentioning, however, that at least one study has used Google Analytics software, using seven standard GA reports, to identify potential design problems and to improve the design and content of a library website [3]. Fang [3] also recognised the importance of investigating the usability of the site after implementing the suggested improvements. The usage of the site was tracked by GA for the first time (the pre-modification time range) for 22 days, while usage data concerning the site were collected by GA for the second time (post-modification time range) for 22 days. However, Fang [3] used only three metrics provided by GA standard reports (new visitors, return visitors and number of pages viewed during each visit) to investigate the improvements to the design. The values of the metrics showed that the modified design improved the usability of the site: the new visitors increased by 21%, the returning visitors increased by 44%, and the number of people who viewed more than three pages increased by 29%. However, there is a lack of research to illustrate the improvements of the usability of an e-commerce website by considering a matrix of advanced web metrics that provide an overall picture of the usability of a site in terms of different areas (i.e. navigation, content).

#### **4. AIMS AND OBJECTIVES**

The aim of the research described here was to investigate the improvements in usability and performance of an e-commerce website after it had been redesigned and its usability improved.

The specific objectives of the research were:

- To use a matrix of advanced web metrics, calculated by GA software, to measure the usage of an e-commerce website for two time ranges: pre- and post- modification time ranges;
- To make a comparison of the values of the metrics for the two time ranges to evaluate the potential improvements in the usability of the redesigned website.

#### **5. METHODOLOGY**

The research involved an e-commerce case study. It compared the values of advanced web metrics for two time ranges. The first was between September 1 and November 30, 2008, which represents the usage of an e-commerce website before modifications; the second, between June 1 and August 31, 2009, represents the usage of the site after being redesigned and its usability improved.

In order to identify usability problems of the site and to suggest improvements, a user testing method was employed after the usage of the site had been tracked and measured by the web metrics after the first time range.

In order to use GA software to track the usage of the e-commerce website, it was necessary to install the required script on the company website. In order to employ the user testing method, a task scenario was developed for the site. Twenty users were recruited. Data were gathered from each user testing session using screen capture software (Camtasia), with post-test questionnaires. Observation of the users working through the tasks, in addition to taking comments from the users while interacting with each site, was also undertaken.

A matrix of advanced web metrics was used to measure the usage of the site tracked by GA. This matrix was suggested by earlier research [6]. The matrix includes: specific metrics that could, either individually or in combination, identify potential usability problems on an e-commerce website in relation to six areas (navigation, internal search, architecture, content/design, purchasing process, and the customer service) and specific web metrics which can help to provide useful supplementary information about the site's visitors and its financial performance. Table 1 shows the suggested matrix and the combination of metrics.

The user testing method was analysed by examining: performance data, in-session observation notes, notes taken from reviewing the twenty Camtasia sessions, users' comments noted during the test, and quantitative and qualitative data from the post-test questionnaires.

A report on usability problems and recommendations was sent to the company in February, 2009. The report presented the identified usability problems after they had been summarised, categorised and explained. Recommendations and suggestions for each of these problems were also explained and presented for each problem. All the suggestions were accepted by the company and implemented in the new design of the site. Visitors' characteristics (see Section 6.6) were taken into consideration in the new design, which was launched in May, 2009.

Table 1. Web metrics indicating the overall usability of a site

Usability Problem Area	Web Metrics
Navigation	Bounce rate, average number of page views per visit, average searches per visit, percentage of visits using search, percentage of click depth visits.
Internal Search	Average searches per visit, percentage of visits using search, average number of page views per visit, percentage of click depth visits, search results to site exits ratio.
Architecture	Percentage of time spent on visits, average searches per visit, percentage of visits using search, percentage of click depth visits, average number of page views per visit.
Content/Design	Percentage of click depth visits, percentage of time spent visits, bounce rate. <i>Specific pages</i> : top landing pages metrics (bounce rate, entrance searches and entrance keywords), top content pages metrics (bounce rate, average time on page and percentage of site exits), top exit pages metrics (percentage of site exits).
Purchasing Process	Order conversion rate, percentage of time spent visits, cart completion rate, checkout completion rate. <i>Specific pages</i> : cart start rate, checkout start rate and the funnel report.
Customer Service Visitors' Metrics	Information find conversion rate, feedback form conversion rate. Ratio of new to returning visits, visitor engagement index, language, operating systems, browsers, screen colours, screen resolutions, flash versions, Java support, connection speed.
Financial Performance Metrics	Average order value, average revenue per visit, average items per cart.

## 6. RESULTS

This section reviews the results obtained from the advanced web metrics, using the GA method, for the two time ranges. It presents the potential improvements in the usability of the investigated website in terms of seven areas. It also reviews the visitors' characteristics of the site during the investigated time ranges. It is worth mentioning that based on the aim of this paper, only the results obtained from the metrics are presented. Results obtained from the analysis of the user testing method are not presented.

## 6.1 Navigation

The metrics used to investigate the navigational problems of the site indicated that the navigation of the site had been improved. This can be seen clearly by considering the values of:

- The bounce rate metric which decreased compared to the first time range (Appendix, metric 5)
- The average page views per visit which increased from 17 to 20 (Appendix, metric 2) meaning that visitors viewed more pages during their visits.
- Despite the fact that the values for metrics 12 and 13 increased compared to the first time range (Appendix, *average searches per visit* and *percent of visits using search*), the low value of these metrics could suggest that the site had good navigation so that a search facility was not needed or alternatively that there were problems with the search facilities.
- The percentages of medium click depth visits to the site which increased (Appendix, metric 4).

## 6.2 Internal Search

The metrics used to examine the usability of the internal search of a site showed that the usage level of the internal search facilities of the site had improved compared to the first time range (Appendix, metric 12, *average searches per visit* and metric 13, *percent of visits using search*). Furthermore, the value of the metric which provides an indication of the accuracy of the search results decreased compared to the first time range (Appendix, *search results to site exits ratio*). This means that the changes implemented to the internal search facilities of the site were useful and improved the usability of the internal search facilities of the site.

## 6.3 Architecture

The metrics used to investigate the architectural problems of the site indicated that the possibility of having architectural problems on the site had been decreased compared to the first time range. Despite the fact that the usage of the internal search facilities of the site increased (Appendix, metrics 12 and 13), the relatively low rate of usage indicates that the architecture of the site had fewer problems as visitors were able to navigate through the site. Furthermore, the increased percentage of visits with medium click depth for the site, together with the decreased number of visitors who spent little time on the site (i.e. their visits did not exceed 3 minutes in duration) provide evidence that the potential possibility of having architecture problems of the site has been decreased.

## 6.4 Content/Design

The metrics used to examine the content/design problems of a site indicated that the content of the site improved and visitors appeared to be more interested in the site's content compared to the first time range. This was obvious by considering the values of two metrics: the *percentage of high click depth visits* which increased (Appendix, metric 4) and the *percentage of low time spent visits* which decreased (Appendix, metric 3). However, the metrics indicated that most visitors spent less than 3 minutes on the site (Appendix, metric 3). These metrics implied that there were new usability problems with some content on the site.

The value of the *bounce rate* metric decreased, which also indicated that the potential usability problems in the content or design of the site had been reduced (Appendix, metric 5). The metrics of the top ten landing pages, top content pages and top exit pages also showed that the usability of specific pages within the site had improved but identified specific new pages within the site that had possible usability problems. For example, the top ten landing pages of the site included the home page and nine pages illustrating products (six out of the nine pages were included in the top ten landing pages of the first time range). The bounce rate for the six pages decreased which indicated that the usability problems of these pages had been reduced. However, the high bounce rates for the three pages suggested that users were unimpressed with either the content or the design of the pages.

## 6.5 Purchasing Process

Metrics related to the purchasing process showed that usability problems in the overall purchasing process of the site had been reduced. For example, the value of the *order conversion rate* metric (Appendix, metric 9) of the site increased which indicated that visits resulting in an order increased. Also, the value of the *high percentage of time spent visit* metric increased which suggested that more visitors were engaged in a purchasing activity on the site. Furthermore, the increased value of the *cart completion rate* and *checkout completion rate* metrics (Appendix, metrics 17 and 19) suggest that the usability problems concerning the purchasing process of the site had decreased.

A similar issue was found with specific pages that made up the purchasing process. The two purchasing process metrics (*cart start rate* and *checkout start rate*) and the funnel report showed that the usability of specific pages had been improved:

- The increased value of the *cart start rate* metric compared to the first time range (Appendix, metric 16) suggested that usability problems on product pages had been reduced.

- The value of the *checkout start rate* metric also increased compared to the first time range (Appendix, metrics 18), suggesting that the usability of the pages containing the 'go to checkout' button had been improved.

The funnel report also indicated that the usability problems regarding specific pages in the purchasing process of the site had decreased.

## 6.6 Customer Service

The value of the information find conversion rate metric for the twelve customer support pages, which were identified by the owner of the site, was increased. This provides evidence that it was easier to find and visit the customer support pages compared to the first time range. The value of the feedback form conversion rate metric also increased (Appendix, metric 15) which indicated that visitors to the site were interested enough to send feedback to the web master.

## 6.7 Visitors' Characteristics

The results of the eight metrics that described the characteristics of the computers and Internet browsers used by the site's visitors, together with the connection speed of their network during the second time range, were investigated. The results indicated that visitors' characteristics were similar to those of the first time range. The results of the metrics that described the behaviour of visitors to the site showed that the value for the *ratio of new to returning visits* metric decreased compared to the first time range (from 1.54 to 1.15), which means that the number of returning visits was greater than the number of new visits. The value for the *visitor engagement index* metric for the site increased from 1.54 to 2.09 (Appendix, metric 7), which might indicate that more repeat visitors were engaged and therefore came back to the site compared to the first time range.

## 6.8 The Financial Performance of the Site

The results of the metrics that described the site's ability to generate revenue and to cross-sell showed that the financial performance of the site had improved significantly compared to the first time range after the usability of the site had been improved. This can be clearly seen by considering the value of two metrics: the average order value and the average revenue per visit. The value of the average order value metric increased from \$106.20 to \$150.30 (Appendix, metric 8) and the value for the average revenue per visit metric (Appendix, metric 10) increased from \$1.14 to \$6.31. Furthermore, the value of the average items per cart metric increased compared to the first time range from 4 to 6. This means that visitors were more interested in buying more items from the site (Appendix, metric 11).

## 7. CONCLUSION

This research used a matrix of advanced web metrics, calculated by using GA software, to investigate improvements in the usability and performance of an e-commerce website. A comparison of the values of the metrics was undertaken for two ranges of time, before and after changing the design of the site.

The results showed that the matrix of advanced web metrics was a useful tool for providing an overview of a site's usability, in terms of indicating potential usability problems on a site overall and on some specific pages, and for monitoring improvements in a site's usability.

The metrics showed that the web traffic and performance of the site increased significantly after changing the design of the site and improving its usability. The metrics illustrated improvements in the usability of the site in terms of six areas (navigation, internal search, architecture, content/design, customer service and purchasing process), and in the site's financial performance. This emphasises the importance of considering the usability of e-commerce websites in order to improve their success and performance.

The results offer a base for future research. The next step will be to employ user testing and/or heuristic evaluation to identify usability problems on a site based on the issues raised by the advanced metrics of GA software.

## REFERENCES

1. Barnard, L., Wesson. A Trust Model for E-commerce in South Africa. SAICSIT 2004, 2004, 23-32.
2. Chen, S.Y., Macredie, R.D. An Assessment of Usability of Electronic Shopping: a Heuristic Evaluation. J. International Journal of Information Management. 25, 2005, 516-532.
3. Fang W. Using Google Analytics for Improving Library Website Content and Design: A Case Study. J. Library Philosophy and Practice, 2007, 1-17.
4. Freeman, M.B., Hyland, P. Australian Online Supermarket Usability. Technical Report, Decision Systems Lab, University of Wollongong, 2003.
5. Hasan, L., Morris, A. & Proberts, S. Using Google Analytics to Evaluate the Usability of E-Commerce Sites. HCI International 2009 Conference, San Diego, CA, USA, July 19-24, 2009, Proceedings. Lecture Notes in Computer Science 5619 Springer 2009, 697-706.
6. Hasan, L. Usability Evaluation Framework for E-Commerce Websites in Developing Countries, PhD Thesis, *Institutional Repository, Loughborough University*, 2009.
7. Inan, H. Measuring the success of your website: a customer-centric approach to website management (Software Book). Hurol Inan., 2006.
8. Kaushik, A. Web Analytics, an Hour a Day. Wiley Publishing, Inc., 2007.
9. Malacinski, A., Dominick, S. & Hartrick, T. *Measuring web traffic, part1*, <<http://www.ibm.com/developerworks/web/library/wa-mwt1/>>, [accessed 12.01.07], 2001.
10. McFadden, C. Optimizing the online business channel with web analytics. <http://www.webanalyticsassociation.org/en/art/?9>, 2005.
11. Nielsen, J. Usability 101: Introduction to usability. Useit.com. <[useit.com/alertbox/20030825.html](http://useit.com/alertbox/20030825.html)>. 2003.
12. Nielsen, J., Norman D. Web-Site Usability: Usability on The Web Isn't A Luxury. Information Week, <http://www.informationweek.com/773/web.htm>, 2000.
13. Nielsen, J. & Mack, R. L. Usability inspection methods (Eds.). John Wiley & Sons, New York. 1994.
14. Peterson, E. Web analytics demystified. Celilo Group Media and CafePress, 2004.
15. Phippen, A., Sheppard, L. & Furnell, S. A practical evaluation of web analytics. Internet Research, 14(4), 2004, 284-293.
16. Tilson, R., Dong, J., Martin, S., Kieke E. Factors and Principles Affecting the Usability of Four E-commerce Sites. 4th Conference on Human Factors and the Web (CHFW), AT&TLabs, USA, 1998.
17. Web Analytics Association. <http://www.webanalyticsassociation.org>, [n.d.].



## APPENDIX

### THE VALUES OF ADVANCED METRICS FOR THE TWO TIME RANGES

No.	Metric		Results	
			Time Range 1 (September 1 and November 30, 2008)	Time Range 2 (June 1 and August 31, 2009)
1	Average Time on Site		00:06:52	00:10:50
2	Average Page Views per Visit		17.00	20.00
3	Percentage of Time Spent Visits	Percentage of low time spent visits (between 0 seconds and 3 minutes)	60.16%	48.76%
		Percentage of medium time spent visits (between 3 and 10 minutes)	21.67%	28.45%
		Percentage of high time spent visits (more than 10 minutes)	18.17%	22.79%
4	Percentage of Click Depth Visits	Percentage of low click depth visits (two pages or fewer)	31.29%	23.15%
		Percentage of medium click depth visits (between 3 to the value of metric 2)	42.57%	48.33%
		Percentage of high click depth visits (more than the value of metric 2)	26.14%	28.52%
5	Bounce Rate for All Pages		22.77%	15.16%
6	Ratio of New to Returning Visits		1.54	1.15
7	Visitor Engagement Index		1.54	2.09
8	Average Order Value		\$106.20	\$150.30
9	Order Conversion Rate (OCR)		1.07%	3.07%
10	Average Revenue per Visit		\$1.14	\$6.31
11	Average Items per Cart Completed		4	6
12	Average Searches per Visit		0.07 (product search) 0.01 (advanced search)	1.2 (product search) 0.8 (advanced search)
13	Percent Visits Using Search		2.14% (product search) 0.20% (advanced search)	5.17% (product search) 1.4% (advanced search)
14	Search Results to Site Exits Ratio		0.79 (product search) 0.60 (advanced search)	0.32 (product search) 0.28 (advanced search)
15	Feedback Form Conversion Rate		11.69%	12.4%
16	Cart Start Rate		5.94%	10.15%
17	Cart Completion Rate		18.07%	24.23%
18	Checkout Start Rate		3.63%	5.18%
19	Checkout Completion Rate		29.55%	35.66%
20	Ratio of Checkout Starts to Cart Starts		0.61	0.84

# TAILORABLE FLEXIBILITY: MAKING END-USERS AUTONOMOUS IN THE DESIGN OF ACTIVE INTERFACES

Federico Cabitza, Iade Gesso and Stefano Corna  
*Dipartimento di Informatica Sistemistica e Comunicazione (DISCo)*  
*Università degli Studi di Milano-Bicocca*  
*Viale Sarca, 336 20126 Milano (Italy)*

## ABSTRACT

In this paper we present the visual editing tools of the WOAD framework that we propose to make end-users autonomous in tailoring their digital documents to their ever changing needs. After a brief outline of the essential points regarding the architecture and conceptual model of WOAD, we describe the tools we have developed to allow end users to both create their own document templates and augment them with proactive and context-aware rules (mechanisms in WOAD).

## KEYWORDS

Active Document, Mechanisms, Visual Editor, EUD, WOAD, ProDoc

## 1. INTRODUCTION

In the past six years we conducted a series of observational studies in the hospital domain (e.g., [Cabitza et al., 2009a]) with the aim to elicit the requirements that clinicians, seen as particularly demanding users of their electronic patient records (EPRs), considered most important to avoid that the digitization of their document tasks would disrupt their daily cooperative work practices. We found clinicians very aware of this risk, recognizing that EPRs can end up requiring more work effort, or imposing organizational and procedural constraints [Nowinski et al., 2007]. The requirements we gathered can be grouped in three main categories, i.e. *support*, *autonomy* and *flexibility*, and are the basis on which we conceived the architecture called WOAD [Cabitza and Gesso, 2011] and realized ProDoc [Cabitza et al., 2009b], a prototypical EPR based on that architecture. The ‘support’ class encompasses traditional data-oriented functionalities that support users in doing their work (e.g., data filtering or document printing). On the other hand, ProDoc is a proof-of-concept application that stresses the classes of ‘autonomy’ and ‘flexibility’. Obviously, these classes are correlated. Flexibility is a necessary requirement to make EPRs, and in general electronic document applications to be used in complex and distributed organizations, tailorable to the needs of the users, especially when requirements can change over a regular basis to support an unpredictable flow of work.

Differently from traditional approaches, we think that users have to make their documents flexible “on their own”, i.e., by being autonomous with respect to ICT specialists and application vendors. Thus, we consider autonomy an important precondition that must be guaranteed to reach actual flexibility. Specifically, WOAD is a computational platform that is aimed at making users autonomous in two distinct activities: (i) building and maintaining their own digital documents (conceived as sets of modular and reusable components), (ii) augmenting documents with simple rules that are executed asynchronously and are triggered by the context and the content that users progressively fill in. This makes WOAD an end-user development environment [Lieberman et al., 2006] where users can interact with two specific visual editors. The former allows users to create the templates of their electronic documents, i.e., the interfaces to their information system, in a manner that mimic what they do with the templates of their paper-based documents with a traditional word processor, i.e., by placing fields and input elements into blank templates and then modifying them in the same way. The second is a visual rule editor that facilitates users in creating rules, i.e. simple if-then constructs that make those documents ‘active’, i.e., able to react to the user interaction and

adapt to the progress of work to support it even more flexibly. Although any kind of rule can be created to mimic corresponding business rules between the data managed at interface level, we intend the editor described in this paper as a tool to tune the interface according to the local conventions currently in use in a specific department [Cabitza et al., 2009a] in order to increase collaboration awareness [Dourish and Bellotti, 1992] among users in the cooperative setting. Through the WOAD rule editor, users can visually create conditions over the templates they have previously created (e.g., with the template editor mentioned above), and couple them with sets of simple operations that can both act on any part of a document and modulate how the document content should look like, conveying what in [Cabitza et al., 2009a] has been called Awareness Promoting Information (API). This is any additional indication at interface level (e.g., message box, background color, additional metadata) that could make users aware of what is going on in their work setting, and help them recall useful knowledge to cope with the situation.

In the following sections, we will outline the WOAD main concepts and describe how end-users can interact with visual editing tools we conceived to make users satisfy their needs for flexibility by their own.

## 2. RELATED WORKS

Relatively few research contributions can be found on visual composition of document templates and interfaces, especially compared to commercial solutions that have reached impressive user-friendliness over time (e.g., SAP NetWeaver Visual Composer, Adobe LiveCycle Designer and FileMaker Pro). Some of these applications present affinities with our solution. For instance, *WebSheets* [Wolber et al., 2002] is a WYSIWYG tool that allows end-users to create dynamic web pages with the capability to access and modify database contents without programming. *WebSheets* is based on the Programming by Example (PBE) technique, and allows to use spreadsheet formulas inside the design environment. Visual form editing is not a prerogative of traditional and web-based applications: in [Chande and Koivisto, 2006] authors present a mobile editing platform that allows to create forms using mobile devices. Using a graphic wizard, users can add new fields, specifying type and name. Subsequently, users can arrange the form UI using a WYSIWYG editor that allows to set the field appearance order only. In [Yamazaki et al., 2000], authors present their *Visual TDL Document Editor* (VTDE) that can be used to create and edit EPR templates. In particular, this editor allows to manage the *Template Definition Language* (TDL), an XML dialect that allows to represent both content and structure of a template, and has been developed to promote template sharing between clinicians and institutions. VTDE is split in two windows: the former is a text editor and the latter displays the graphic representation of the editing results of text editor. Thus, the only visual aid provided to users is the possibility of checking ‘on-the-fly’ how the template will appear while users typeset the related TDL code in the text editor. Moreover, differently from our solution, VTDE lacks in supporting the definition of rules for the dynamic change of templates (described using comments) that must be implemented manually.

On the other hand, due to the diffusion of ontologies in many domains, visual rule editors are becoming more and more common; yet these are not usually associated to the interface templates where users fill in and consult the data that these rules have to match and transform. In this line, Andersen Consulting developed *Eagle* [Davidowitz, 1996], which is a set of tools, architectures and reusable components for the externalization of business-object behaviors. Users can tailor these behaviors specifying rules using the Smalltalk language. Users compose rules using a “point-and-click” editor that allows to visually check the correctness of a rule building and displaying the related syntax tree. On the other hand, users are not aided in any way in creating a rule, which must be composed typing the executable code like in any traditional development environment. In [Chen et al., 2002], authors present a real-time alerting system that has been conceived to act on the database of a clinical information system, according to user-defined alert rules. Users can define their alert rules using a graphic rule editor that lists some physiological parameters and helps them in specifying the conditions through a set of buttons. These allow to define two kinds of conditions, which are respectively called *basic rules* (i.e., a subset of the logic expressions) and *advanced rules* (e.g., time bounds, parameter changes, and any combination of them). Finally, in [Li et al., 2010] it is presented a visual rule editor to create rules for optimizing airplane load planning. This editor allows users to drag and drop the model elements (e.g., condition, action and flow) picking them up from a palette, in order to visually compose the flow of their rules. Conditions are expressed using the Object Constraint Language (OCL) that is a standard language for object relationship description.

### 3. DOCUMENTS IN THE WOAD FRAMEWORK

WOAD encompasses both a *reference software architecture* and a *conceptual model*, and is grounded on the concepts of “web of documental artifacts” and “active document” [Cabitza and Gesso, 2011]. Documents are composed by two independent but strictly intertwined parts: a *passive* part and an *active* part. The former contains the contents that users fill in and arranges them according to the description that is hold into a *template*, and the latter is a set of *mechanisms*, i.e. how we call if-then rules in the WOAD environment.

A *template* defines how the document is structured. In particular, a template defines the *didgets* (i.e., ‘documental widgets’) that compose the structure of a document and how they are topologically arranged inside the document. Didgets are the reusable instances of the *datoms* (i.e., ‘documental atoms’). These are modular data structures that represent specific aspects of the reality of interest, and encompass a coherent set of data fields and the definition of the features of these fields in terms of both simple constraints (e.g., data type, format or mandatoryness) and graphic styles. In particular, a datom is defined using a XML dialect (i.e., the *XForms* syntax), which is used both for the data model (i.e., the data structure and the related constraints) and UI (e.g., a multiple choice field can be rendered using either a list of values or a set of checkboxes). A datom also embeds indications about the visual aspect of its fields (i.e., *Cascading Style Sheet* (CSS) syntax).

Templates are stored into a XML data structure that mainly stores the references to the didgets that have been used within a template. As mentioned above, each didget reference is characterized by topological information (i.e., position and size) and a set of attributes that can modify either its scope (e.g., the globality attribute, see Section 4) or appearance. Moreover, this data structure allows to store the changes of an existing template in a not destructive way, adopting a chronological versioning system: each new template version is labelled with its creation timestamp. On the other hand, the document instances and their related didget contents are stored into other distinct data structures.

*Mechanisms* are rules that makes WOAD documents “active” and proactive with respect to their content. Mechanisms can be defined at level of either i) datom, if they are conceived to be valid in general and across different resources (e.g., an email address validation); ii) didget, i.e., when they refer to didgets that are in one or more specific document templates; iii) specific didgets for specific resources, e.g., when a doctor wants to activate a reminder triggered by the blood pressure value of John Doe only. Mechanisms are composed by some conditions that are defined over the didget content (*antecedent* or *if-part*), and simple actions (*consequent* or *then-part*) that are executed whenever all the conditions are met. Mechanisms are triggered by human interaction with documents and modulated according to the content of didgets. Any application behavior can be associated to the mechanism consequent, if a programming interface is available. Mechanisms are classified according to how they act on the document content, and it is possible to distinguish between mechanisms that (i) modify the content, e.g., to edit or correct values in data fields; (ii) modify content attributes and metadata, e.g., timestamps, status flags, urgency attributes; (iii) use the content, e.g., print (parts of) it or check its quality; (iv) transmit the content from one system to another, e.g., through an email; (v) change the appearance of content, e.g., modifying the background color or the font family; (vi) route documents and build flows of work, e.g., allowing users to link a document to another or subordinate fill-in operations on certain (portions of) documents to the same operations on other (portions of) documents.

Mechanisms execution can be seen as a process of API generation, i.e., any operation that modifies the affordance and appearance of documents and their content, and possibly conveys to the user additional information (e.g., a message) that makes her aware about some condition in the context of document use.

### 4. THE TEMPLATE EDITOR

The *Template Editor* (TE) is a prototypical application that is based on the *Oryx Editor*<sup>1</sup> (see [Decker et al., 2008]), a web-based editor conceived to model business processes. Oryx Editor is based on a plug-in architecture that allows to extend it easily by adding new visual editing features (e.g., Petri Nets).

According to the arrangement of the Oryx Editor user interface, the TE is split in three areas (see Figure 1). The left area contains the palettes with the lists of existing datoms (B in Figure 1), which have been

<sup>1</sup> *Oryx* is an open source project, which has been developed by the *Business Process Technology* research group of the *Hasso-Plattner-Institute* (<http://bpt.hpi.uni-potsdam.de/Oryx/>).

created with a datom editor that we will not describe for brevity's sake, and didgets (C), i.e. datoms already placed in some template, with the addition of a list of some support elements (e.g., text and image placeholders) that are commonly used also in standard documents (A). The central area (D) represents the document (with the proportions of the A4 paper sheet) in which the users can put the didgets that the document must contain. Finally, the right area (E) contains a list of properties that are related to the currently selected element in the central area, according to its type (e.g., the URI and the alternative text of an image).

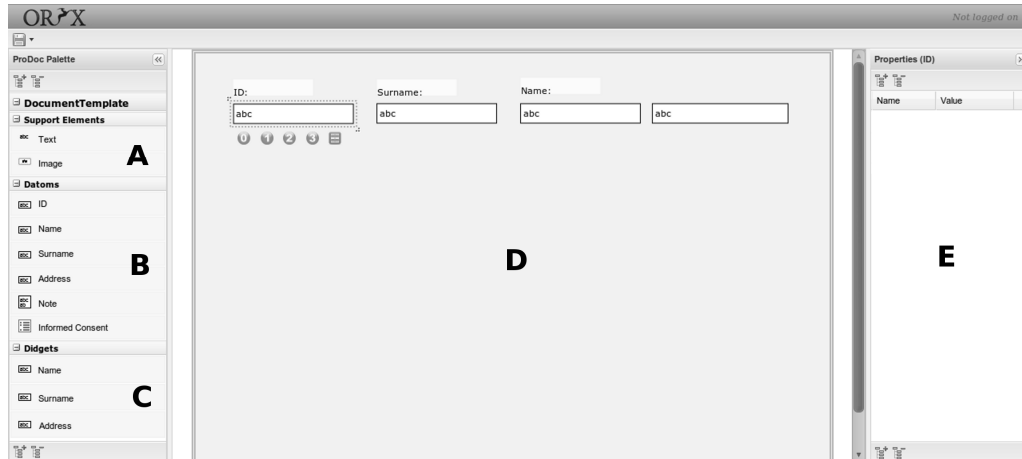


Figure 1. The Template Editor UI.

A user who wants either to create or edit a template has simply to select datoms from the Datoms palette (in the left column), drag them over the central area and drop them at the desired position (see D in Figure 1). Once a datom has been dropped, a new didget is created into the Didget palette (C). Also already existing didgets can be placed into the template, and this can be done simply dragging and dropping them; in this latter case, the drop operation does not create any new didget, but prompts user in specifying the level of globality of the didget (see below).

Grouping didgets into a separate palette with respect to datoms allows to make didgets available outside the template in which they were created, and this allows to reuse them into other templates. Reusing didgets allows for sharing contents between both different documents, either based on the same template or not, and different resources (e.g., all the patients of a hospital ward). Sharing content policies are defined according to the didgets *globality level* that can assume four different values (see Table 1). Users can set the desired level of globality simply selecting the corresponding icon into a graphic menu that appears directly under the graphic representation of the currently selected didget (see Figure 1). When users drop a datom and create the related new didget, the latter will hold only local data (e.g., the value of the daily measurement of the patient's temperature that practitioners inscribe on the *Daily Sheet*). Acting on the above mentioned menu, users can set the didgets to share their content between either all the instances of a document that is based on a specific template and related to a specific resource, instances of documents that are based on different templates and related to a specific resource (e.g., some portions of a patient's personal data), or all the document instances both based on any template and related to any resource.

Users can also specify if the fields of a didget must be displayed only once ("single didget") or have to be repeated for a certain number of times ("multiple didget") in the document. For instance, this can be useful to handle those data that is needed to organize in tabular format (e.g., the set of vital parameters of a newborn within few moments from delivery). In this case, the structure of the didget (described by the related datom) is used to define both the format and the arrangement of the table rows.

This setting can be applied through the same graphic menu that can be used to set the globality level (see the last icon in Figure 1). Once a didget has been declared as "multiple didget", users can set the right number of repetitions through a didget property that can be found in the list of properties in the right area of the editor.

Making the users autonomous in building their own documents in a what-you-see-is-what-you-get manner allows for increasing timeliness, flexibility and 'tailorability' [Ardito et al., 2009] with respect to both creating and modifying operations, according to the changes of the local needs. An example can be the

need to add a new field to a document (e.g., a checkbox) that allows the clinicians to indicate if the patient's informed consent has expressed or not. This new field imposes the constraint that all operations on a portion of the document has to be inhibited if the checkbox has not been marked.

Table 1. The globality levels of a didiget

	Data Shared Between		
	Instances	Templates	Resources
<i>G0</i>	✗	✗	✗
<i>G1</i>	✓	✗	✗
<i>G2</i>	✓	✓	✗
<i>G3</i>	✓	✓	✓

In a traditional information system, to address this need requires the involvement of software analysts and developers with the aim of applying a set of modifications that could concern the whole system, with the possibility of having to wait for a substantial amount of time.

On the other hand, through the TE, users can quickly add any new feature to the documents themselves, without to involve any other professionals. They have just to edit the document template, picking up the "Informed Consent" datom, drag and drop it at the desired place in the template. Similarly, using the *Mechanism Editor* (see Section 5 for more details), users can also be autonomous in adding the application logic that prevents from or enables the editing of the fields in the same document.

## 5. THE MECHANISM EDITOR

The *Mechanism Editor* (ME) is the tool we designed to allow users to create and edit mechanisms in a visual and step-like manner. Like the TE described in Section 4, also the ME is a visual editor based on the Oryx Editor. ME provides users with a user-friendly GUI that allows them to compose mechanisms that pertain to the documents defined with the TE. With this tool, users can define mechanisms by means of drag & drop operations and selections from short closed-option menus, so to make rule composition easy also for users with little or no experience in declarative programming. The GUI of the editor is horizontally split into three areas (see Figure 2).

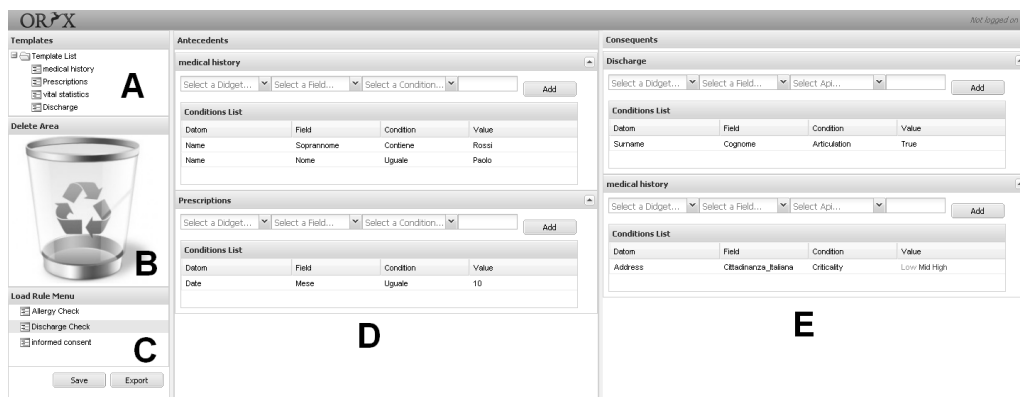


Figure 2. The Mechanism Editor user interface.

The left area is in turn divided in three sections. Top section (A in Figure 2) contains the list of all the existing document templates previously created with TE. Templates can be dropped both into central (D) and right (E) areas to build respectively the conditions (if-part) and the actions (then-part) of the mechanism that pertains to those particular templates. The bottom section of left area (C in Figure 2) is the load mechanism menu and contains the list of all the previously saved mechanisms and two controls (buttons) by which to save and export a mechanism. An existing mechanism can be loaded and displayed in the main area just by a double click on the mechanism item in this list. The left area also contains a trash area (B in Figure 2) in which users can drop any action, condition or mechanism that they want to delete.

The typical interaction with the tool will be explained through a running example, taking the case of the informed consent checkbox (see Section 4) as exemplificatory of the main passages and operations performed by the user. In this case, the user needs to create a mechanism that inhibits the fill in operations on a document where a specific field, namely the informed consent checkbox, has not yet been checked. The user starts the composition by picking up the document template from the list in the left area (A in Figure 2) and dropping it into the central area (D in Figure 2). The central area of the ME is the area in which the if-part of a mechanism is composed. The if-part of a mechanism contains the conditions that the system must match to the content of the document. Those conditions can be defined on one or more templates (and their didgets), as well as on the basis of some environmental variable (e.g., the system date and time, or the current user).

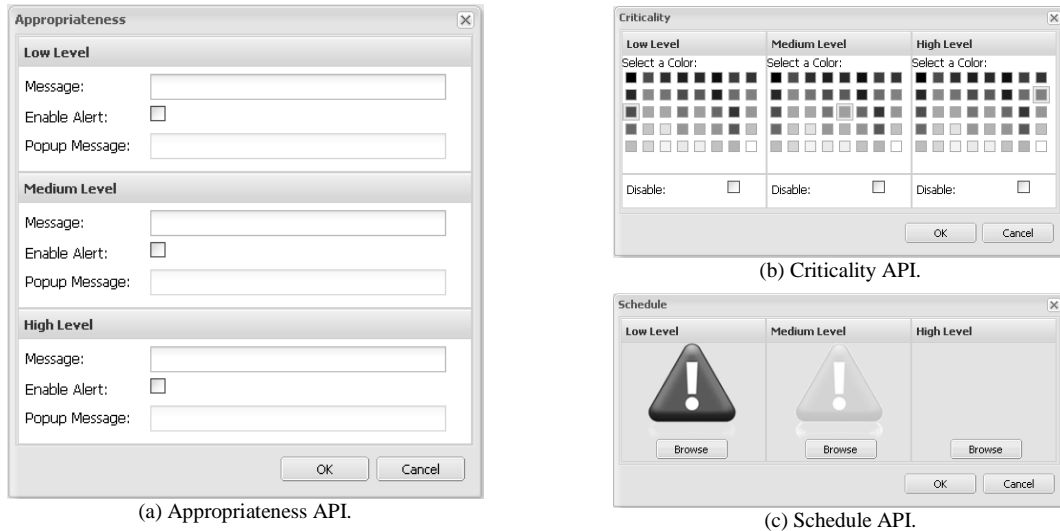


Figure 3. Examples of API additional property forms.

At this step, the user can compose the necessary conditions through the interface (see Figure 4) that is displayed in the central area; she selects the informed consent datum from the datum dropdown, and the informed consent checkbox from the field dropdown. Then she selects the “equals” constraint from the third dropdown, she writes in the textbox the “unchecked” value and pushes the “Add” button to complete the condition. Once the condition has been created, the user can start to compose the then-part of the rule. The then-part contains the actions to be triggered when all the defined conditions in the left part have been met. To this aim, the user drops the previously chosen template in the right area and composes the action through the interface that is displayed afterwards in that area. The user selects the datum that contains all the fields that she needs to protect from the datum dropdown and selects the “inhibits” API from the API dropdown. The “inhibits” API has been rendered so as to make all the selected datum fields read-only (and look like they cannot be edited) and to change their background color to light red. Finally, the user pushes the “Add” button to complete the action.

In the then-part, users can associate various kinds of controls: in the current prototype we have focused on a kind of functionality that is mainly aimed at facilitating collaboration awareness and at conveying an API that fits the current work context. Each type of API, if triggered, produces a different effect on the document; therefore, we have defined specific API-related parameters to be specified in the then-part. When the user pushes the “Add” button, the ME displays a property window that contains a form with the selected API parameters. In Figure 3 are shown three examples of API property window, respectively the *Appropriateness* properties (a), the *Criticality* properties (b) and the *Schedule* properties (c).

Before delving into those details, we need to introduce the concept of *activation level of a mechanism*. In this context, with *level* we mean the degree of expertise or familiarity with the system that users possess in managing their documents. The level of expertise is a parameter that the system can calculate according to different factors. In the analysis we undertook in the hospital domain, we considered three different levels of expertise: namely *low*, *medium* and *high*; to this respect, low-, medium- and high-expertise users in the medical domain can be respectively novices, residents and any senior nurse or doctors that have to interact

with the system. User expertise level is an important aspect to consider when defining a mechanisms because we want to avoid that a mechanism, when triggered, conveys to the user information that is either redundant, overloading, difficult or just meaningless according to her level of expertise and familiarity with the system. For those reasons, the ME allows to define the behavior of a rule for each of these levels of expertise. To allow this parameterization, each API property form is split in three identical sections, one for each expertise level.

The *Appropriateness* form (AF) (Figure 3a) contains the parameters to customize the Appropriateness API. Its aim is to provide indications on what could be appropriate to do or not regarding the document content through the generation of different kinds of inline messages. The AF allows the user to define the text associated with those messages. The AF is horizontally split in three sections, each of these sections contains two text fields and a checkbox. The first text field holds the text message that the API will display directly through the electronic document. The checkbox allows the user to enable the generation of an additional pop-up message, which is defined through the second text field. Toggling the checkbox respectively enables or disables the second text field.

Datum	Field	Condition	Value
Name	Name	Equals	Luke
Address	City	Contain	Milan

Figure 4. The Mechanism Editor condition composition form.

The *Criticality* form (CF) (Figure 3b) contains the parameters to customize the Criticality API. Criticality API indicates the need to consider the situation reported in the document as critical. This API, when triggered, changes the color style of fields or document sections. The CF allows the user to choose the colors she prefers to use in the document rendering. The CF is split vertically in three sections, each section contains a color palette and a checkbox. The color palettes allow user to choose the color for the specific expertise level mentioned above. The checkbox (unchecked by default), when checked, disables the effect of this API for a specific expertise level (e.g., in a clinical scenario, to hide a low criticality situation to expert users to not cause information overload).

Finally, the *Schedule* form (SF) (Figure 3c) contains the parameters to customize the Schedule API. This kind of API is conveyed to make users aware of the need to perform tasks that have been previously scheduled or expected on the basis of a timing convention, e.g., a doctor must examine a blood sample within twenty hours from the collection; the users that consult the *Exam Sheet* (ES) during this time span can be reminded of this threshold and after the time limit has elapsed the system generates a warning message according to this mechanism. The Schedule API, when triggered by a mechanism, displays different icons near the fields, on the basis of the user's expertise level, e.g., a clock icon near the expiration date field and on the basis of what the users has selected in phase of mechanism composition. The SF is split vertically in three sections. Each section contains a "Browse" button, by which the user can select an image icon from his desktop, and an area where a preview of the selected icon is shown.

Once the mechanism has been defined, user can save it into the local repository (for future modifications) by pushing the "Save" button in the left area, and then she can convert it into a specific rule-based language (i.e., the Drools DRL) by pushing the "Export" button.

## 6. CONCLUSIONS AND FUTURE WORK

The paper has briefly outlined two tools conceived within the WOAD architecture, focusing on the level of the user interface and interaction. The former of these, the template editor, is a WYSIWYG editor of electronic forms that allows users to define their digital documents in a flexible and modular way; the second tool, the rule editor, complements the former one and allows users to define simple behaviors that are associated to the documents and that enrich their structure and content (e.g., in terms of different affordances)



according to the context and the level of user expertise. In other words, we focused on two prototypical applications that allow users to define both the passive structure and the active behaviors of their interfaces, as well as to maintain them over time in a visual manner.

These two functionalities have been experimented in the context of ProDoc, a prototypical and innovative EPR that adopts a case-oriented stance on the management of clinical documentation; yet these functionalities can be adopted in a much wider class of applications, especially when these are designed to support collaborative work settings and satisfy coordinative needs mainly through documents themselves. To this general aim, WOAD compliant applications are conceived to convey a particular kind of additional information through the user interface in order to promote “collaboration awareness” [Dourish and Bellotti, 1992], i.e., Awareness Promoting Information. This is done according to simple rules that a visual editor allows to create over the structural components of documental interfaces, so to facilitate end-users in tweaking their tools even if they lack specific programming skills. This places our research program within the scope of both the End-User Development and Interaction Design fields. Consequently, our future work will focus on how to further improve the usability of the tools presented in this paper and to make them more adaptive to the level of technical skill and domain expertise exhibited by their users. To this aim, the empirical work that inspired the conception of the WOAD framework and its proof-of-concept application, ProDoc, will continue to validate its applicability in other domains where we have gained an initial positive feedback [Locatelli et al., 2010]. In this line, our future work will include the design of a wider set of experiments aimed at validate the efficiency of our solution with respect to other similar applications.

## REFERENCES

- Ardito, C., Lanzilotti, R., Mussio, P., Parasiliti Provenza, L., and Piccinno, A. (2009). Redefining the Roles of Users and Designers in Interactive System Lifecycle. In *CHIItaly 2009, June 17-19, Rome, Italy*.
- Cabitz, F. and Gesso, I. (2011). Web of Active Documents: An Architecture for Flexible Electronic Patient Records. *Biomedical Engineering Systems and Technologies, LNCS: 127*, pages 44–56. Springer Berlin Heidelberg.
- Cabitz, F., Simone, C., and Sarini, M. (2009a). Leveraging coordinative conventions to promote collaboration awareness. *Computer Supported Cooperative Work (CSCW)*, 18:301–330.
- Cabitz, F., Simone, C., and Zorzato, G. (2009b). ProDoc: an electronic patient record to foster process-oriented practices. In *ECSW'09. Vienna, Austria, September 9-11, 2009*, pages 119–138. Springer.
- Chande, S. and Koivisto, A. (2006). Mobile form-editor: a push based group communication tool. In *Mobility '06, New York, NY, USA. ACM*.
- Chen, H.-T., Ma, W.-C., and Liou, D.-M. (2002). Design and implementation of a real-time clinical alerting system for intensive care unit. In *AMIA Symposium*, page 131.
- Davidowitz, P. (1996). Externalizing Business-Object Behavior: A Point-and Click Rule Editor. *The Smalltalk Report*, 6(1):4–10.
- Decker, G., Overdick, H., and Weske, M. (2008). Oryx – An Open Modeling Platform for the BPM Community. *Business Process Management, LNCS: 5240*, pages 382–385. Springer Berlin/Heidelberg.
- Dourish, P. and Bellotti, V. (1992). Awareness and coordination in shared workspaces. In *CSCW'92*, pages 107–114, New York, NY, USA. ACM Press.
- Li, F., Tian, C., Zhang, H., and Kelley, W. (2010). Rule-based optimization approach for airline load planning system. *Procedia Computer Science*, 1(1):1455–1463.
- Lieberman, H., Paternò, F., Klann, M., and Wulf, V. (2006). End-User Development: An Emerging Paradigm. In *End User Development*, volume 9, chapter 1, pages 1–8. Kluwer Academic Publishers.
- Locatelli, M., Viviana, A., and Cabitz, F. (2010). Supporting learning by doing in archaeology with active process maps. In *eLearning 2010, Freiburg, Germany. 26 - 29 July 2010*.
- Nowinski, C., Becker, S., Reynolds, K., Beaumont, J., Caprinia, C., Hahn, E., Peresa, A., and Arnold, B. (2007). The impact of converting to an electronic health record on organizational culture and quality improvement. *International Journal of Medical Informatics*, 76(1):S174–S183.
- Wolber, D., Su, Y., and Chiang, Y. T. (2002). Designing dynamic web pages and persistence in the WYSIWYG interface. *IUI '02*, pages 228–229, New York, NY, USA. ACM.
- Yamazaki, S., Takabayashi, K., Hirose, Y., Satomura, Y., Kamiyama, T., and Matsuo, H. (2000). Visual Editor for Template Design of Electronic Patient Record. In *AMIA Symposium*, page 1160.

# PERSONIFICATION OF TOPICS WITH CONVERSATIONAL AGENTS

Jean-Paul Sansonnet<sup>1</sup>, François Bouchet<sup>2</sup> and William A. Turner<sup>1</sup>

<sup>1</sup>LIMSI-CNRS, BP 133 F-91403 Orsay cedex France

<sup>2</sup>SMART-Lab McGill University, 3700 McTavish, Montreal, QC H3A 1Y2 Canada

## ABSTRACT

Nowadays, ordinary people can freely access a large amount of information through Internet. Then issues of production and transport of computational entities are superseded by new requirements like: enticement, acceptability and understanding. To achieve these goals, we propose an approach based on dialogical interaction between novice users and conversational agents achieving the mediation of computational entities. The originality of the approach is based on the principle of personification of an entity by an agent, where the agent and the entity share a single identity.

## KEYWORDS

Computational objects, Conversational agents, Personification.

## 1. INTRODUCTION

### 1.1 Facilitation of Ordinary People Interaction with Abstract Topics

Following the development of the Internet culture, general public can now access a plethora of computational entities through Web applications and services. Those entities, which we'll refer to as **topics** in this article, are data structures synthetized in order to mediate information (technical, encyclopedic, statistical...) to a targeted audience, with a given goal (didactic, institutional, social, entertaining, advertising...). Because of their abundance, these topics are in competition in order to: 1) reach their audience: it's the *enticement* issue (Galon, 1999), 2) not to be discarded immediately: it's the ergonomic *acceptability* issue (Davis 1989), 3) taken seriously: it's the *believability* issue, 4) to be properly understood: it's the *communication goal* issue. Hence, there is a need to develop tools for facilitating ordinary people interaction with such abstract topics, where the facilitation can be measured in terms of the four four listed above.

In this paper, we propose an approach to facilitate ordinary people interaction with abstract topics, which is based on their personification by Conversational Agents. Ongoing research in the very active research field of conversational agents (Maes 1994) has shown that they can increase the user-friendliness (Lester et al. 1997) useful for enticement), the acceptability and the believability (Hayes-Roth 2004) of the system, as well as increase the performance of the human-agent couple to learn new information. Hence, they appear as a promising approach to address issues regarding interaction of general audiences with topics.

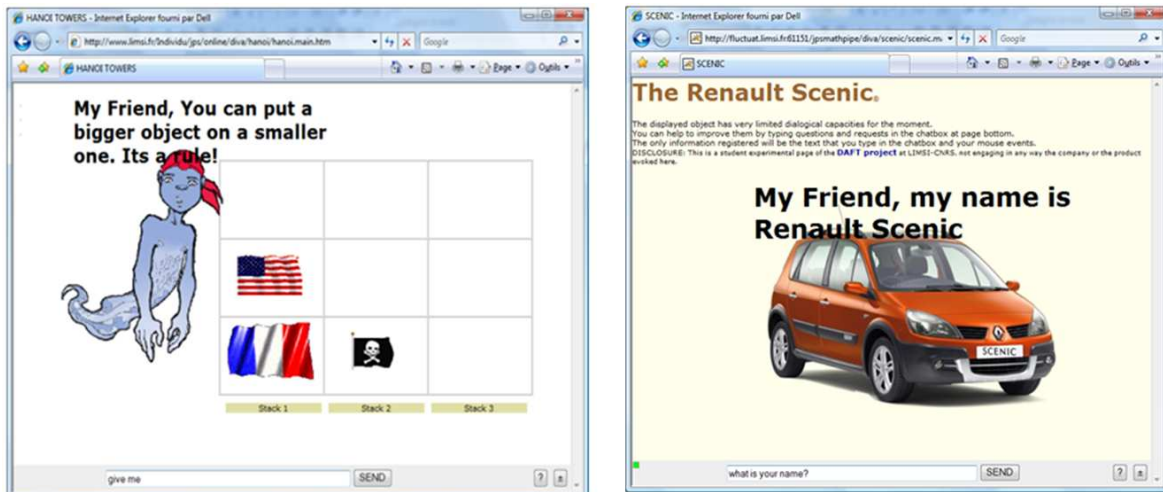


Figure 1. Left) A typical assistant agent integrated into a DIVA web page: the user interacts in natural language with the agent (in this case, it is a cartoon-like animated character but DIVA provides also realistic ones) about a simple game, which acts here as a topic — right) In this conversational situation, the user interacts directly with the topic (it is a commercial product personified by a static picture), which answers users' questions *as if the car were a person*. The DIVA toolkit is freely available for research and education purposes at <http://www.limsi.fr/~jps/online/diva/divahome/index.html>

## 1.2 An Architecture for the Personification of Abstract Topics

Traditionally, conversational agents are personified through virtual graphic characters and are studied in the context of Embodied Conversational Agents (ECA) (Cassell et al. 2000). In such a case, agents represent human-like entities and attempt at expressing relevant, believable, human cognition and psychology. Human cognition is often based on rational reasoning approaches, such as classical rule-based systems (ACT-R, SOAR) or now popular BDI-agent technology for Intelligent Agents (Rao and Georgeff, 1991). More recently, there have been attempts at implementing human psychology phenomena into Intelligent Agents, like for example CoJACK (Norling & Ritter 2004) (Evertsz et al. 2008) for the BDI platform JACK (Howden et al. 2001). Following the pioneering works on Believable Agents (Rousseau and Hayes-Roth, 1996), authors such as (Malatesta et al. 2007) use personality traits to create different expressions of behaviors, especially by influencing the appraisal part of the OCC theory (Ortony et al. 1988) etc.

When we are concerned with the personification of non human-like entities like the topics defined above, especially abstract ones, we face a contradiction: 1) They are not human-like in essence and their embodiment (e.g. their appearance on screen) is not human-like; 2) We state in principle that people should interact with them “as if they were human-like entities”.

Taken in its generality, this issue has been discussed from the philosophical point of view (Nagel 1974). It has also been studied by psychologists since Piaget in the '30s and more recently studies upon children have linked personification of non human-like entities with notions such as *animism* prompting definitions: “*Since Piaget [], psychologists have often claimed that young children are animistic and personifying []. Animistic means labeling inanimate objects as living as well as attributing characteristics of animate objects (typically humans) to inanimate objects and making predictions or explanations about inanimate objects on the basis of knowledge about animate objects. Personification means the extension of human attributes to any nonhumans*” (Inagaki et al. 1987).

In order to support these requirements, we have to design a software architecture for conversational agents that enables the integration of conversational topics into Web pages. This architecture will be based on previous works on Conversational Assistant Agents (Sansonnet 1999, Sansonnet et al. 2002) and more precisely on the DIVA toolkit. Actually DIVA stands for DOM-Integrated Virtual Agents, which emphasizes the unique feature of DIVA agents to interact in natural language with users and to control directly the inner structure of the content of the Web page they assist (Bouchet and Sansonnet 2007, Xuetao et al. 2011). Figure 1-left shows a typical DIVA assisted Web page. The DIVA architecture has been easily adapted to

support conversational interaction with non human-like entities, as shown in Figure 1-right. This raises the issue of the *interactional content* between users and topics, which is discussed through a case study developed in the next section.

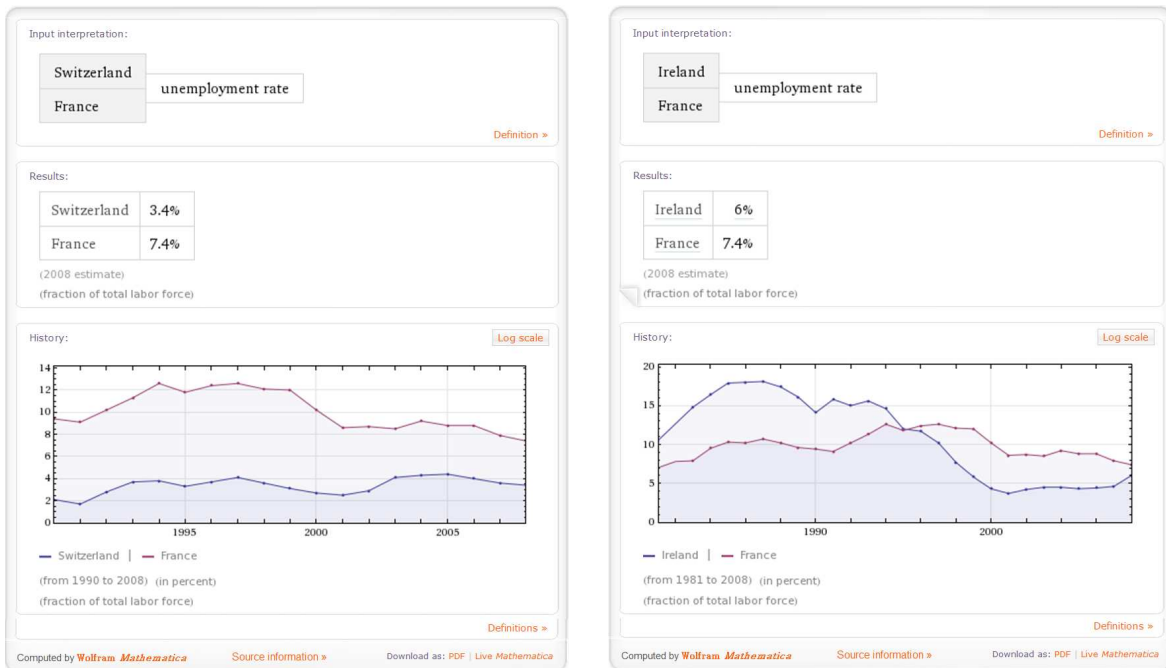


Figure 2. Left) Alpha answer to: "unemployment rate switzerland france" (upper part); right) Alpha answer to: "unemployment rate ireland france" (upper part).

## 2. PERSONIFICATION PROCESS OF AN ABSTRACT TOPIC

In this section, we develop a case study so as to exhibit the issues raised by the conversational interaction between ordinary people and abstract topics, on the basis of the metaphor of (Inagaki et al. 1987) mentioned above. The outline of the case study is as follows: first we choose a typical abstract subject, synthesized from Web sources that can be relevant for many people browsing the Internet. In the next sub section, a model of the agent associated with the chosen topic is briefly discussed and in further sub sections, we propose excerpts of both rational and psychological interactions.

### 2.1 Choice of a Topic to Personify

The case we propose here is willingly related to a very abstract phenomenon of interest: « Unemployment in France for the last 20 years ». This topic has been mentioned by many books and documentary films, and websites about it<sup>1</sup> provide data (including a lot of graphs) and text comments. Regardless of the stability of the information (*i.e.* static website versus wikis) they provide, data is difficult to process automatically for creating an associated agent. However, recent advances in Semantic Web for extracting data in order to build new documents have been particularly promising:

*Alpha* by Wolfram Research<sup>2</sup>, takes a Question&Answer (Voorhees 2002) type of request and provides as a result, not a mere text but a full Web page related to the user's question. Moreover, data retrieved by Alpha is now accessible directly from the latest version of Mathematica (8.0) under a symbolic format.

<sup>1</sup> See <http://en.wikipedia.org/wiki/Unemployment> for a first list of references

<sup>2</sup> <http://www.wolframalpha.com>

*QWiki*<sup>3</sup> uses the Alpha technology to build even more sophisticated Web pages, both in terms of content and presentation.

Those two examples show that in the future, users will indeed be facing an abundance of topics of interest automatically synthesized as answers to their(s) request(s), and this is the reason why we have also decided to use the results of Alpha requests regarding our aforementioned topic of interest. Two successive requests provide results shown in Figure 2 (only the upper part of the results Alpha Web page is given here).

## 2.2 Building the Agent Model

In section 1.2 we have mentioned that authors have tried to integrate psychological phenomena into traditional Intelligent Agents architectures. However, psychologists have shown how rational and psychological human behaviors are extremely entangled (Frijda 2006) (Scherer et al. 2001). In this context, we have proposed in previous works (Bouchet and Sansonnet 2009) a framework, called R&B — standing for Rational and Behavioral agents. This framework dedicated to the experimentation of conversational scenarios, where both rational and psychological behaviors are mixed according to various strategies.

In this section, we will assume, as it has been explained in section 2.1, that the agent and the topic are merged into a single entity, meaning that the user talks to the topic as if it were a person. Hence, one must be able to provide a conversational agent tightly related to the topic, which must be a multimodal mediator of the topic informational content. Based on framework R&B principles, two modes of answer are then required:

*Rational expression:* users should be able to put natural language questions to the agent-topic and to receive factual answers in return, which requires storing facts into the Knowledge Base (Kbase) of the agent.

*Psychological expression:* the agent and the topic being viewed as a unique entity, this identity must be reflected in the mental states of the agent and it should affect its multimodal expression. To achieve this, facts should be related to the mental states of the agent, *i.e.* to the part of its Kbase that models its psychology.

The association between objective facts and their rational and psychological expression is not neutral: it requires decisions and interpretations from the agent's designer:

- For the rational part, it is about selecting which facts should be entered into the model or not, *i.e.* what notions the agent will be able to comment and talk about.

- For the psychological part, it is about defining influence operators over the mental states of the agent.

In both cases, one must have some kind of judgment regarding the data provided in the answer, which will depend on *who* will express the information (that is who is supposed to be the agent?) and to *whom* it will be transmitted (that is who is the user?). Hence it is crucial to properly define the roles and the communicative goals of the actors of the interactional session.

For instance, in Figure 2-Left, one can read the objective fact  $\phi =$  “*unemployment rate in France is now of 7.4% against 3.4% in Switzerland, and the difference is stable over 20 years as shown by the graph*”. With regard to this kind of information, the agent's designer has to adopt a role among the following ones: Swiss, French or Other. Independently from the difference between the numerical values, the designer also needs to have a communicative goal: *e.g.* to enjoy about the stated differences (whether they are positive or negative), to bemoan it or to minimize their relevance. In the same way, the user also has a role (he/she can be Swiss, French or Other) and has a communicative goal that can be obtained either explicitly or inferred by the agent according to the user's profile. This simple combinatory entails a set of possibilities that can't be all considered here – indeed, if we consider only the possible nationalities and if we use a stereotypical model of nationalism, we could have:

User	Agent	Agent's reaction about $\phi$
French	French	sad
	Swiss	arrogant
Swiss	French	proud
	Swiss	enjoying

That's the reason why in the rest of this case study we'll assume both the agent and the user are French, and the communicative goal is to emphasize the absolute value of the difference between the countries, and to regret it if it's unfavorable. Table 1. provides symbolic information related to rational and psychological

<sup>3</sup> <http://www.qwiki.com>

expressions for this topic. Globally, the rational part stores the absolute value of the difference (Rate-unempl-F) and the difference between the two involved countries (Dif-unempl-FS, Dif-unempl-FI), while in the psychological part, two moods are connected to the difference between the countries according to formulas (1) and (2) in Table 1.

Table 1. Summary of the symbolic model of unemployment rate difference between France and Ireland/Switzerland.

Domain	Operator type	Attribute	Format	Link
Rational	$\delta_i$	Rate-uneempl-F	List	plot values
		Dif-unempl-FS	List	plot values
		Dif-unempl-FI	List	plot values
Psycho. $\Delta_{mood}$	Satisfaction	scalar [-1,1]	F - I	(1)
		Confidence	scalar [-1, 1]	F (S+I) / 2

## 2.3 Dialogical Interaction using the Rational Expression

Let's assume the user asks from the agent a purely factual question about the topic:

USER: "What's the current unemployment rate in France?"

Moreover, let's assume that among the patterns associated to the rational attribute Rate-unempl-F is the rule:

QUEST & /unemployment rate france/  $\Rightarrow$  Comment(Rate-unempl-F)

Triggering this rule means the agent reads the value of the attribute Rate-unempl-F in its KBase and can comment it in a multimodal way to the user. The commenting strategy is linked to attributes with rules like:

Comment (Rate-unempl-F) :=

**IF** val < 0. **THEN** Play SORRY & Say "(Sadly | Unfortunately), it's @val %" ;

**ELSE** Play HAPPY & Say "The unemployment [rate] in France is only of @val %"

Where @val is dynamically replaced by the value of the attribute in the answer.

This interaction is above all rational : the user asks for a value and gets it in return. It is however possible to add to the answer some psychological content through a judgment over the raw fact. We can use for this two expressive modalities:

Operator Say builds a natural language sentence (using options [] or picking among alternatives for variety with |, etc.) and displays it in the answer area.

Operator Play triggers an animation that emulates the expression of emotions or emotes<sup>4</sup>.

For instance, we could have the following multimodal answer ("sadly" willingly emphasized here):

AGENT: "Sadly, it's 7.4%" + SAD

## 2.4 Dialogical Interaction using the Psychological Expression

In this section, we provide an example involving the static traits and the dynamic moods composing the mental states of the agent. We analyze the different possible combinations of the following elements:

*Static traits* are specified once and for all by the agent's designer depending on: a) the personality profile associated to the topic, b) the type of interaction expected (information, entertainment...). Their automatic elicitation is beyond the scope of this study.

*Dynamic moods* are linked once and for all by the agent's designer to the topic variables, *i.e.* data that is evolving with time, either because a) the topic itself is dynamic (its values can change throughout the session) or b) the topic is static (like here) but contains values that have been evolving in time (cf. the graphs

<sup>4</sup> In traditional ECA environments, simple emotions (joy, anger, surprise, etc.) or more complex so-called emotes gestures (beware danger!, so bored, etc.) are expressed through graphical animations of the agent's character. In this case, the graphical representation being more abstract (*e.g.* plot, technical graphics) we have to find ways to express emotions through alternative modalities. These ways are not discussed here because their elicitation, in each specific context, is an issue in its own. See for example (Sansonnnet et al. 2010) in the context of agents in ambient environments.

illustrated in Figure 2), in which case it should be possible to prompt the user “to browse through time” in order to get back to case a.

**Example: Linking mental states to data**

We will consider one trait and two moods:

Binary (yes/no) trait Talkative which means that: yes = the agent provides lots of information; no = the agent restricts to minimal information when asked.

Moods Satisfaction et Confidence, which take values in [-1, 1], 0 being neutral. Satisfaction is related to cognitive happiness while confidence indicates the degree of certainty about a fact.

When it is possible to access data used to generate the graphs, one can link the values of the graphs to the internal mental states, as shown with formulas (1) and (2) in Table 1. More complex relationships can take into account the derivative or integral over a period of time. The graph from Figure 2-right, using formula (1) shows that the psychological attribute Satisfaction of the agent evolves over time from a negative to a positive value when the two curves cross in 1995 (thus triggering a data event), then gets back to a neutral value in 2000: the agent *psychologically lives* the evolution. The same analysis can be done for Confidence with the formula (2).

In this context, if the user asks the agent how it feels, like one would do for a companion:

USER: “How do you feel?”

Using a rule of the list of patterns associated to the psychological behavior of the agent, like:

QUEST & /(you feel|are you)/  $\Rightarrow$  Express(Talkative, Satisfaction, Confidence)

Would lead to a set of possible reactions according to the static (Talkative) and dynamic (Satisfaction and Confidence) mental states of the agent, summarized in the heuristic below:

Express(Talkative, Satisfaction, Confidence) :=

Talkative	Satisfaction	Confidence	Natural language answer	Emote
Yes	>0	>0	“I’m [quite very] confident”	HAPPY
		<0	“I’m (ok fine)”	NEUTRAL
	<0	>0	“I’m sad, but determined”	NEUTRAL
No	>0	<0	“Ouch! Everything is wrong!”	SAD
		>0	“Everything is fine”	NEUTRAL
	<0	<0	“(Err Hmm)...”	NEUTRAL
	<0	*	“...”	A LOOF

## 2.5 Discussion

### Extension: argumentative interactions using additional agents

In the proposed approach, the topic of interest interacts directly with the users. A more traditional approach is also possible, where additional agents can be added on the Web page in order to interact with users **about** the topic, as shown in Figure 1-left. This can prove useful to extend the architecture with *argumentative* agents.

Indeed, it is a current trend, particularly observed in online social networks, to share opinions and to debate in a participative context. A way to motivate the general public to get interested in a given topic and also to better understand it, would then consist in offering a personification of this topic with an *argumentative agent*, which extends the concept of assistant conversational agent by providing it with argumentation skills (Kakas & Moraitis 2002) (Bentahar et al. 2007), and by making it have positive/negative/neutral opinions about phenomena related to the topic. In this case study, we had considered that the situation was stable with regard to roles and communicative goals but it can also be dynamic. Moreover this can be applied with additional agents:

— with a single conversational agent: for instance, at the beginning of the session the agent has a neutral position and changes its attitude depending on the user’s identified opinion, either to take the same one (empathetic interaction) or an opposed one (entering into debate).

— with three conversational agents: in this case, each of them takes one of the possible positions (positive, negative, neutral) in order to enter into an argumentative interaction with the user.

In all cases, the objective is to induce users to interact with the facts associated to a phenomenon of interest through an activity that they consider enjoyable in itself (*e.g.* a chatting activity).

#### **Generalization: towards automatic personification**

In this paper we have chosen to present the process of personification of an abstract entity through a single case study, which raises the issue of the genericity of our approach. Beyond the mere feasibility demonstrated here, we can put forward two main arguments in favor of genericity:

1) A lot of works have been carried out in automatic extraction of information from data (data mining etc.). Hence, it is possible to rely on existing algorithms for detection of status differences (Figure 2-left) or “data events” (*e.g.* curves crossing in Figure 2-right) to automatically fill rule patterns as shown in Table 1.

2) Conversational roles are quite generic (*e.g.* to take a positive, negative, neutral position about a status or an event) as well as communicative goals. They can be captured into rule-based heuristics that in turn can be instantiated for each particular domain.

#### **Experimentation: evaluation of the impact of conversational topics**

Because this work is preliminary, it is premature to state that conversational topics have a better impact (in terms of the four main factors listed in section 1.1, that is enticement, acceptability, believability and understanding) upon ordinary people than non interactive topics like those presented and discussed in blogs for examples. However these factors are well-studied in ECA community and it has been proved that conversational agents can bring significant improvements. Consequently, we can rely on these results and confidently expect that they will extend to non human-like entities.

Moreover, implementations of conversational situations involving non human-like entities have already been developed and can be accessed on the DIVA Web page (URL at Figure 1.). They will be an operational support for further investigations upon the four factors.

### **3. CONCLUSION**

We have presented an approach meant to facilitate the mediation between novice users and the abstract computational entities, such as those produced automatically in the Web. This approach is based on the principle of personification, which identifies the subject of interest (the topic) with a conversational agent, given both rational and psychological reasoning skills. Using our previous works on assistant conversational agents, as well as on works in psychology, we have proposed a method to build an agent associated to an entity, using a case study that illustrates the various problems encountered during this process.

The case study led here provides a first positive answer to the question of feasibility. It also shows that it's a vast domain still largely to explore, involving many different disciplines (agents, natural language understanding, psychology, semantic web, data extraction...). Two main questions would then have to be addressed: how to move toward an automation of the synthesis of the agent-topic identity once given a particular topic (and also how to profile both the designer's and the users' communicative goals), and what experiments could confirm the validity of with regard to enticement, acceptability, believability and understanding.

### **REFERENCES**

- Bentahar J. et al. 2007. Using Argumentative Agents to Manage Communities of Web Services. In *Proceedings of the 21st International Conference on Advanced Information Networking and Applications Workshops - Volume 02*. AINAW '07. Washington, DC, USA: IEEE Computer Society, p. 588–593.
- Bouchet F., Sansonnet J. P. 2007. Caractérisation de Requêtes d'Assistance à partir de corpus, *MFI'07*, Paris.
- Bouchet F., Sansonnet J. P. 2009. A framework for modeling the relationships between the rational and behavioral reactions of assisting conversational agents, *Seventh European Workshop on Multi-Agent Systems (EUMAS 09)* Ayia Napa, Cyprus.
- Cassell J. et al. ed. 2000. *Embodied Conversational Agents*, MIT Press.
- Davis F. D. 1989. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 13(3), p.319-340.
- Ellsworth P. C., Scherer K. R. 2003. Appraisal processes in emotion. In R. J. Davidson, K. R. Scherer, & H. H.



- Goldsmith, éd. *Handbook of affective sciences*. Series in affective science. New York, NY, USA: Oxford University Press, p. 572-595.
- Evertsz R. et al. 2008. Realistic Behaviour Variation in a BDI-based Cognitive Architecture. In *Proc. of SimTecT'08*. SimTecT '08. Melbourne, Australia.
- Frijda N. H. 2006. *The Laws of Emotion*, Psychology Press.
- Galon D. 1999. *The Savvy Way to Successful Website Promotion : Secrets of Successful Websites, Attracting On-Line Traffic, the Most Up to Date Guide to Top Positioning on Search Engines*. Trafford on Demand Pub.
- Hayes-Roth B. 2004. What makes characters seem life-like? In H. Prendinger & M. Ishizuka, éd. *Life-like characters: tools, affective functions, and application*. Springer, p. 447-462.
- Howden N. et al. 2001. JACK intelligent agents-summary of an agent infrastructure. In *Proc. of the 5th International Conference on Autonomous Agents*.
- Inagaki K., Hatano G. 1987. Young Children's Spontaneous Personification as Analogy. In *Blackwell Publishing for the Society for Research in Child Development*, p. 1013-1020.
- Kakas A. C., Moraitis P. 2002. Argumentative Agent Deliberation, Roles and Context. *Electronic Notes in Theoretical Computer Science*, 70(5), p.39-53.
- Lazarus R. S. 1966. *Psychological stress and the coping process*, New York, NY, USA: McGraw-Hill.
- Lester J. C. et al. 1997. The persona effect: affective impact of animated pedagogical agents. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. SIGCHI. Atlanta, GA, USA: ACM, p. 359-366.
- Maes P. 1994. Agents that reduce work and information overload. *Commun. ACM*, 37(7), p.30-40.
- Malatesta, L. et al. 2007. Agent personality traits in virtual environments based on appraisal theory predictions. In *Artificial and Ambient Intelligence, Language, Speech and Gesture for Expressive Characters*. AISB 07. Newcastle, United Kingdom.
- Nagel T. 1974. What it is to be a bat? *The philosophical review*, 83(4) p. 435-450.
- Norling E., Ritter F. E. 2004. Towards Supporting Psychologically Plausible Variability in Agent-Based Human Modelling. Dans *Proceedings of AAMAS'03*.
- Ortony A., Clore G. L., Collins A. 1988. *The Cognitive Structure of Emotions* Cambridge University Press., Cambridge, UK.
- Rao A. S., Georgeff M. P. 1991. Modelling Rational Agents within a BDI Architecture. Dans R. Fikes & E. Sandewall, éd. *Proceedings of Knowledge Representation and Reasoning*. KR&R-91. San Mateo, CA, USA: Morgan Kaufmann, p. 473-484.
- Rousseau D., Hayes-Roth B. 1996. *Personality in Synthetic Agents*, Knowledge Systems Laboratory: Stanford University.
- Sansonnet J. P., Sabouret N., Pitel G. 2002. An Agent Design and Query Language dedicated to Natural Language Interaction, *AAMAS'02*, Bologna, Italy.
- Sansonnet J. P., Bellik Y. 2010. Expressing agents personality through ambient modalities, *WACA 2010*, Lille p. 47-48.
- Sansonnet J. P. 1999. Description scientifique du projet InterViews - The InterViews project, *LIMSI Report*.
- Scherer, K.R., Schorr, A. & Johnstone, T., 2001. *Appraisal processes in emotion: Theory, methods, research*, Oxford University Press.
- Voorhees E. M., 2002. Overview of the TREC 2002 Question Answering Track. In *Proc. of the Eleventh Text REtrieval Conference (TREC)*. Gaithersburg, Maryland, USA, p. 115-123.
- Xuetao M., Sansonnet J. P., Bouchet F. 2011. Définition d'un agent conversationnel assistant d'applications Internet à partir d'un corpus de requêtes, *Revue TSI*, Vol. 29 (10), p.1123-1154.

# EVALUATION OF GESTURAL INTERACTION WITH AND WITHOUT VOICE COMMANDS

Joana Neca and Carlos Duarte

*LaSIGE, University of Lisbon*

*FCUL, Edifício C6, Campo Grande, 1749-016 Lisboa, Portugal*

## ABSTRACT

Gesture interaction has been gaining traction. Although mostly used in smaller devices, like smart phones and tablets, new devices, like Microsoft's Kinect, open a new range of interaction possibilities for interaction with larger surfaces. In this paper we explore how users interact with large surfaces when they are not in touching distance from them. We studied two interaction scenarios. In the first scenario users had only the possibility to interact through gestures. In the second, users were able to combine gestures and speech commands. The paper discusses how users employ gestures differently when speech commands are available and what types of actions are more adequate in each of the scenarios.

## KEYWORDS

Gestures, Speech, Multimodalities, Evaluation, Design.

## 1. INTRODUCTION

Recent movies and literature have been popularizing gesture interaction in 3D work environments. Outside the fiction realm, we have been witnessing an increase in interaction scenarios where alternatives to using peripheral devices as a means to interact with the system have become available. Our body is now a viable interaction mechanism. As a result, it is vital to understand advantages and disadvantages of gestural interaction, in an effort to make it more efficient and to take advantage of a natural human communication medium [McNeill98].

Most studies on the use of gestures so far have focused on virtual games, where the body is used as an input device [Silva09]. Benefits from using our body as an interaction device are a better spatial understanding [Schuchardt07] and a heightened sense of awareness [Usoh99]. [Kwon05] combined the use of body sensors with motion tracking cameras and, through usability evaluations, concluded that users find gesture based interfaces to be more attractive and become more focused in completing their tasks. Most of the works where gestural interaction occurs in 3D environments, like a 3D gallery of museum objects [Stenger09] or direct manipulation of objects in a 3D map environment [Yin10] are focused on technical aspects. Studies on how users interact in these scenarios are still missing, and they will play an essential role in promoting an even widespread adoption of gestural interfaces.

In this paper we present a study on gestural interaction with and without speech commands, applied to large interaction surfaces, where the user interacts at a distance, thus having no physical contact with the projected surface. With this study we tried to understand the different gestures that people make in two scenarios that distinguish themselves by offering users, or not, the possibility of complementing gestures with voice interaction. To that end, we asked participants in the study to manipulate objects or the working area in two applications. With this, we tried to achieve the following:

1. Understand how users take advantage of gestural interaction, with and without support from voice commands;
2. Which actions are more adequate to be performed by gestural interaction with or without voice commands;
3. Which are the most appropriate, comfortable and intuitive gestures and speech commands for the different actions.

## 2. STUDY FRAMEWORK AND PREPARATION

The main goal of this study was to compare how gestural interaction is used with and without support from speech commands, when there is no physical contact between the user and the interaction surface. We hope to better understand, for a set of actions relevant to this scenario, how to select the best interaction scenario, derive guidelines for designing such applications, and contribute to build interaction dictionaries for both gestures and speech commands by collecting what the users really employ naturally when interacting.

To this end, we selected two applications that could benefit from gestural interaction. For each application we identified a set of actions that are relevant in each of the interaction scenarios that are going to be studied, gestures with and without support from speech commands. We then grouped the actions into meaningful tasks. During task execution, test participants were completely free to select which gestures and which speech commands to use.

### 2.1 Scenarios and Actions

Two flash applications (Figure 1) for image manipulation have been selected for the experiments. Some of the actions to be evaluated are common to both applications, while others are specific. In the first application, a table with images, the images are randomly distributed over a surface, and it is possible to move the images in a 3D space (bring the images closer or further from the surface). In the second application the images are spread out across a rounded wall, which can be navigated, while the images themselves can be zoomed in. Both applications, given their spatial nature, can benefit from gestural interaction, with or without voice commands, particularly when the interaction space is being projected out of reach of the user.



Figure 1. Applications used in the experiments. (1) Table with images; (2) Wall with images

For each application the following set of actions has been selected: 1) Table with images: zoom (in and out), rotate (clockwise and counterclockwise), move, overlap images, delete, undo and redo; 2) Wall with images: zoom image (in and out), show next/previous image, rotate wall (left and right), zoom wall (in and out). In order to exercise these actions, tasks have been defined that imply manipulating images in a meaningful workflow. For instance, instead of asking the user to rotate an image, we placed an inverted image on top of the table and asked the user to examine a detail in the image. This meant that users had to rotate and zoom in on the image, allowing us to understand how the individual actions are used in the context a natural workflow.

### 2.2 Test Participants

A total of 10 individuals, with an average age of 22 years old, took part in the experiment. Eight participants reported having more than 10 years' experience with computers. All reported knowing at least one device supporting gestural interaction (Apple iPhone, iPad, Nintendo Wii, several other smartphones), but only two used them regularly. Eight participants reporting being aware of at least one device supporting voice interaction (screen readers, voice enabled smartphones, in car navigation system), although just one reported using such devices regularly. No participants reported knowing any system integrating both modalities.

## 2.3 Test Procedure

The study began with an explanation of its goals, after which the participants filled a profiling questionnaire. During the experiment, participants had complete freedom to select the gestures and voice commands to use. Each participant interacted with both applications. It was asked of participants to perform tasks, in a way that all actions being studied were executed. Thus all participants had to create gestures and voice commands for all actions. There was no time limit. After completing the tasks, the test participants filled a satisfaction questionnaire where both interaction modes were classified in 5 point scale. It was also asked of participants to try and justify their choices of gestures and speech commands. All the procedure was captured in a video recording. Both the order of presentation of the applications (table and wall with images) and the interaction scenarios (gestures with and without speech commands) were randomized.

## 2.4 Technical Setup

The participants' arms and hand movement was tracked with a Microsoft Kinect. The interaction surface was projected onto a wall, occupying an area of 280 cm per 200 cm (resolution of interaction surface was 1366 per 768 pixel). Participants interacted while standing up, and had no interaction device available. The evaluator remained apart from the participants, out of their field of view, but with excellent visibility to both participant and projection surface. The Kinect allowed participants to control the cursor position. Given that participants had complete control over the creation of the gestures and the speech commands, no recognition was used. Instead, the evaluator controlled the application using keyboard shortcuts, while interpreting the gestures and commands employed by the participants. Participants were not aware of this "limitation" and believed to be interacting with a fully implemented system during the trial.

## 3. STUDY RESULTS

We will present the results in three sections. First we will present and discuss results from the gesture only interaction scenario. Secondly we will present results from the gestures combined with speech commands interaction scenario. Lastly we will compare the results from both scenarios.

### 3.1 Gesture Only Interaction

The first noticeable result is that for some actions most participants use a similar gesture, while for others there is a great variability in the gestures used by participants. For our analysis, we decided that a gesture would be considered standard across the trial population when at least half of the participants would arrive at the same or similar gesture. Tables 1 and 2 present, for each application, how many participants used the standard gesture for the actions where a standard gesture was found. For the ones where no such gesture was found we present the different gestures performed by the participants. Additionally all gestures are textually described.

In the textual description presented in tables 1 and 2 we used the concept of "point". A point is an abstraction of what the participants used to point at the projection surface: some pointed with their hands and arms stretched, others with the hand closed, while others pointed with a finger. As we were unable to find any semantic meaning for the shapes the participants made with their hands we opted to consider all these the same gestures, with only what is being pointed at considered relevant. For illustration purposes Table 3 presents several ways participants made the zoom in gesture, where the common pattern is the two hands being close together by the end of the movement.

In the following paragraphs we will analyze relevant aspects of how participants interacted with both applications and what actions they performed. Additionally, we will also report on how the interaction patterns evolved as the session progressed.

For the delete, rotate image and move the wall actions we failed to find a standard gesture. For the delete action the different gestures performed by participants are completely unrelated. The rotate image and move

the wall actions are more closely related, with the main distinguishing characteristics being the number of points used by the participants. For all the other actions participants converged to a standard gesture.





Table 1. Standard gestures for the table of images

Action	Participants	Gesture description
Zooming	6	Two points parallel to the projection surface moving away (zoom in) or coming together (zoom out)
Overlap	6	One point moving away from the interaction surface
Move	9	One point moving parallel to the projection surface
Stop movir	9	One motionless point
Rotate	3	One point centered at the object and another describing the rotation movement beginning at the first point
Rotate	3	One single point describing the rotation movement (typically with a smaller radius than the alternative above)
Delete	3	Drawing an X shape, with one or two points
Delete	2	Quickly moving a point left and right
Delete	2	Clapping once (two points coming together parallel to the projection surface
Undo & Redo	4	One point in left to right (undo) or right to left (redo) movement parallel to the projection surface

Table 2. Standard gestures for the wall of images

Action	Participants	Gesture description
Zooming on an image	8	Two points parallel to the projection surface moving away (zoom in) or coming together (zoom out)
Previous & Next	9	One point in left to right (next) or right to left (previous) movement parallel to the projection surface
Zooming on the wall	8	Two points parallel to the projection surface moving away (zoom in) or coming together (zoom out)
Move wall	5	Two points moving left or right parallel to the projection surface
Move wall	4	One point moving left or right parallel to the projection surface

Table 3. Different ways to zoom in

Fingertips		All fingers apart	
Fingers together		Fingers in a triangular shape	

As can also be observed in Tables 1 and 2 there are actions for which participants performed the same gestures. Zooming in or out an image is done in the same way as zooming in or out the wall of images. Zooming out is similar to one of the gestures made to delete an image. Undo, redo, previous, next and moving the wall left or right are actions where participants performed exactly the same gestures. This means that it is important to find a way to distinguish between what is the target of the action (image or wall, for instance) or even what is the action intended.

During the session most of the actions were executed more than once. In the following discussion we focus on the zooming actions, but we have observed similar behaviors for other actions. We noticed that as the session went on more participants used what in the end was the standard gesture. This means that gestures made by participants change during usage as participants get more familiar with all the interaction setup. The first time a zooming action was performed a total of 6 participants made the standard gesture presented in Table 1. By the end of the trial 8 participants were doing the standard gesture. It is important to stress out once again that never during the trial were given instructions to the participants regarding what gestures they should perform. This evolution can perhaps be explained by users at the beginning being more prone to experiment, while after using the system for a while they start using it more intuitively, thus converging to what is now probably becoming a universal gesture for zooming actions.

Interestingly the size of the object that is being zoomed seemed to play a role also. Participants converged quicker to the standard gestures when zooming out, and when applying the zoom operations (both in and out) to the whole wall instead of a single picture. A zoomed in image or the whole wall are large objects, while a zoomed out image is a smaller object. For the bigger objects participants used more quickly both hands to perform the gesture, while for the smaller object participants initially revealed some concerns in selecting the object with one of the hands, while performing the action with the other.

Even though both applications offered a 3D space, only one of the actions converged to a standard gesture where depth is an important factor in interpreting the gesture. This is the overlap action, where images had to be brought forward or backward. The majority of participants move their hands closer or further from the intersection surface in order to perform the gesture.

### 3.2 Gestural Interaction with Speech Commands

In a different stage of the trial, participants were requested to perform similar tasks, but were required to use gestures together with speech commands. Tables 4 and 5 present the standard commands for this scenario, as well as the number of participants who used them.

Table 4. Standard commands for the table of images

Action	Command	Participants
Zooming	“Zoom in” “Zoom out”	5
Move	“Drag”	3
Move	“Move”	3
Stop moving	“Stop”	3
Overlap	“Forward”	3
Overlap	“Come forward”	3
Rotate	“Rotate”	5
Delete	“Delete”	6
Undo	“Undo”	5
Redo	“Redo”	5

One of the most relevant analysis in this scenario is understanding how participants employed the gestures when they have the possibility to use speech also. Tables 6 and 7 show how many participants performed a gesture transmitting information regarding the action, or that use the possibly to interact through gestures to be able to point at the target of the action.

When comparing the gestures performed in both scenarios we can observe that some gestures changed. For instance, when using both modalities the 3 different gestures for the delete action are no longer used. Instead, only one participant still makes a gesture meaning delete (two lateral movements). When trying to move the wall more participants make the swiping gesture with only one hand instead of two hands.

One of the more important decision participants made (even if unconsciously) is whether to use the gesture to transmit the action to perform or to just point at the target of the action. The actions which do not have a clear correspondence with any gesture are typically issued by speech. Examples of these actions are delete, undo and redo, which appropriately are actions which participants could not convert to a standard gesture.

Zooming an image or the wall of images is handled differently by the participants. When zooming images most participants use the gesture simply to point (and communicate the action through speech). When zooming the wall most participants use the gesture to communicate the action. The difference between both actions is the need to distinguish the object being zoomed when zooming an image, even when the image is completely zoomed in that it occupies almost the whole projection surface. When zooming the wall participants probably feel no need to specify the object and assume the system interprets correctly the lack of target for the action as applying to the whole interaction setting.

By observing Tables 6 and 7 we may conclude that participants which use the gesture to transmit the action are repeating the same information that is being provided through the speech command. The exception to this are rotate and wall moving actions, where not all the information required to perform the action is transmitted through speech. Participants speak "rotate" and use a gesture to signal the direction the image should rotate. When speaking "rotate" (the wall) participants use the gesture to indicate the direction it should move.

Table 5. Standard commands for the wall of images

Action	Command	Participants
Zoom in	“Enlarge”	4
Zoom in	“Zoom in”	2
Zoom in	“Open”	2
Zoom out	“Close”	6
Next	“Next”	9 <sup>1</sup>
Previous	“Previous”	6
Zoom in on the wall	“Enlarge”	3
Zoom in on the wall	“Zoom in”	3
Zoom in on the wall	“Enlarge all”	2
Zoom out on the wall	“Reduce”	3
Zoom out on the wall	“Zoom out”	3
Zoom out on the wall	“Reduce all”	2
Move the wall to the right	“Rotate”	4
Move the wall to the right	“Rotate right”	3
Move the wall to the left	“Rotate”	4
Move the wall to the left	“Rotate left”	3

Table 6. Pointing and gesturing in the table of images

Action	Pointing	Gesturing
Zooming	7	3
Moving	6	4
Stop moving	6	4
Overlap	8	2
Rotate	2	8
Delete	9	1
Undo & Redo	9	1

Table 7. Pointing and gesturing in the wall of images

Action	Pointing	Gesturing
Zooming	8	2
Previous & Next	1	9
Zooming on the wall	2	8
Moving the wall	0	10

Almost all speech commands were accompanied by a pointing gesture. This means that the speech command was mostly used to identify the action, while the gesture is simply used to provide an argument for the action (the object it applies to). Sometimes the object of the action is also part of the speech command. In the table of images one participant indicated the object for the zooming, moving and overlapping actions. In the wall of images, a total of 3 participants exhibited this behavior. However, this was only applied to zooming actions. In one of the instances, the participant did use speech to select only the object, while the command was issued through a gesture.

### 3.3 Comparing the Two Interaction Scenarios

The satisfaction questionnaires filled at the end of the session allowed a comparison of how participants feel the two interaction modes are adequate to the actions that they had to perform. Table 8 shows the average classification of interaction modes for each action. The scores range from 1 to 5, where 5 means that the interaction is perfectly natural and intuitive, while 1 represents the opposite.

When both gestures and speech were available, overlapping, undo and redo actions were classified higher. This is in accordance to the difficulties experienced by participants in creating a gesture to represent these actions. This difficulty vanishes as soon as speech commands are available.

Although in the wall of images gestures only seems to be preferred, only one action was preferred clearly in this scenario: wall moving. This may be due to that action comprising all images that are rendered on screen at a given moment, thus encompassing the whole interaction area instead of a single object.

We also tried to more fully understand what actions are more difficult to represent with gestures. To that end, we reviewed carefully the videos of the sessions, and timed how long it took participants to come up

<sup>1</sup> In reality these 9 correspond to 3 participants who said “Next” in English, 3 more who said it in their native tongue, and 3 more who said a synonymous of “next” in their native tongue.

with each gesture. Table 9 presents the average time it took participants to come up with a gesture in both interaction scenarios.

Table 8. Average satisfaction (from 1 to 5) in each interaction scenario

	Action	Gestures only	Gestures and speech
Table of images	Zooming	4.8	4.8
	Moving	4.8	4.7
	Stop Moving	4.4	4.2
	Overlap	3.3	3.8
	Rotate	4.6	4.5
	Delete	4.2	4.8
	Undo & Redo	3.1	4.0
Wall of images	Zooming	4.7	4.2
	Previous & Next	4.7	4.4
	Zooming on the wall	4.6	4.3
	Moving the wall	4.8	3.9

Table 9. Average reflection time (in seconds) for each action in each interaction scenario

	Action	Gestures only	Gestures and speech
Table of images	Zoom in	1.3	0.8
	Zoom out	0.2	0.2
	Moving	1.3	1.1
	Stop Moving	0.5	0.4
	Overlap	3.0	3.1
	Rotate clockwise	1.1	1.4
	Rotate counterclockwise	0.2	0.5
	Delete	7.1	0.4
	Undo	7.0	3.2
	Redo	0.6	1.0
Wall of images	Zoom in	2.0	1.0
	Zoom out	0.3	0.1
	Next	0.6	1.2
	Previous	0.3	0.1
	Zoom in on the wall	1.5	1.4
	Zoom out on the wall	0.3	0.1
	Move the wall to the right	0.7	2.1
Move the wall to the left	0.1	0.1	

We can observe that actions that are performed after a similar, but “opposite” action has been performed (like zooming out an image which in the study only was required after a previous zooming in), are quicker than the “opposite” action. This is intuitive, given that most times the participant would use the same gesture, only performing it in the opposite direction.

The action which required more reflection time was the undo action: 7 seconds in gestures only mode and 3.2 with the help of speech commands (which is a reduction in half). Several actions have a significant difference in the time participants took to find a gesture in both interaction scenarios. In the gestures only scenario, delete takes participants 7.1 seconds to find a gesture, while taking only 0.4 seconds when speech commands are available. The explanation is similar to the one for undo and redo actions: it is quite difficult to come up with a gesture to represent these more “abstract” actions, while it is completely intuitive to do it using speech commands. This is also the reason why 9 participants choose to use only gestures to point when they had to perform these actions with speech commands available.

### 3.4 Participant Feedback

Participants provided a variety of justifications for creating the gestures they did. For the zoom in gestures they compare it to the gestures done for opening objects, for example, a door. For both zooming gestures they cited inspiration in movies and the gestures they perform on their smart phones. For zooming and moving the wall of images movies are once again inspiration sources (particularly "Minority Report"). Rotating is



compared to holding a driving wheel or a door knob. Moving an image is compared to grabbing and holding the image while moving. Stopping the movement is likened to dropping the object in order for it to stop. Delete was explained as throwing the image off screen, deleting with a rubber or tearing a paper sheet. Previous and next gestures are inspired by similar gestures in their smart phones or in turning the pages of a book. Some participants failed to find an explanation for the gestures performed, while most reported difficulties in creating gestures for the delete and overlapping actions because they are not used to having those gestures in their devices.

These comments support our findings that gestural interaction, with or without speech commands, becomes more natural, and converges to similar choices by different persons, after using the system for a period of time, and that people also bring the interaction habits from the devices they own previously to new interaction scenarios.

## 4. CONCLUSION

In this paper we have described an experiment to compare the usage of gestures with and without support of speech commands. The results show that users exhibit different patterns of gestures use when in the presence or absence of speech commands. When speech commands are not available, the gestures have to provide information about the action to perform and the target of the action (when applicable). When speech commands are available, the majority of the actions are transmitted through speech while gestures are used to point at the target of the action. Although this is the standard observable use of gestures, it is particularly relevant for more abstract actions, like “undo” and “redo”. For those actions, speech commands are extremely valuable, as can be seen by the difficulties that test participants had to come up with a gesture to represent such actions. This is observable not only by the time it took participants to create such gesture, but also in the variability of gestures designed for those actions. The paper additionally presents the gestures and speech commands our test participants created during task execution and that will be the basis of our future work, where a complete system, integrating those results will be evaluated.

## ACKNOWLEDGEMENTS

This work has been financed by National Funds through FCT – Fundação para a Ciência e a Tecnologia in the scope of project PTDC/EIA-EIA/105062/2008.

## REFERENCES

- Kwon, D.Y. and Gross, M., 2005. Combining body sensors and visual sensors for motion training. *Proceedings of ACE 2005*. Valencia, Spain, pp. 94-101.
- McNeill, D., 1998. Speech and gesture integration. *The nature and functions of gesture in children's communication. New directions for child development*, No. 79, pp. 11-27.
- Schuchardt, P. and Bowman, D., 2007. The Benefits of Immersion for Spatial Understanding of Complex Underground Cave Systems. *Proceedings of VRST 2007*. Newport Beach, USA, pp. 121-124.
- Silva M. and Bowman D., 2009. Body-based Interaction for Desktop Games. *Proceedings of CHI'09*. Boston, USA, pp. 4249-4254.
- Stenger B. and Woodley T., 2009. A Vision-based System for Display Interaction. *Proceedings of British HCI 2009*. Cambridge, UK, pp. 163-168.
- Usoh, M. et al, 1999. Walking > Walking-in-Place > Flying, in Virtual Environments. *Proceedings of SIGGRAPH 99*. Los Angeles, USA, pp. 359-364.
- Yin Y. and Davis R., 2010. Toward Natural Interaction in the Real World: Real-time Gesture Recognition. *Proceedings of ICMI-MLMI'10*. Beijing, China, article 15, 8 pages.

# MULTIMODAL INTERACTION FOR USERS WITH AUTISM IN A 3D EDUCATIONAL ENVIRONMENT

Alessandro Trivilini<sup>1</sup>, Licia Sbattella<sup>2</sup> and Roberto Tedesco<sup>3</sup>

<sup>1</sup>*University of Applied Sciences and Arts of Southern Switzerland, Department of Innovative Technologies(DTI)  
Centro Galleria 2, 6928 Manno, Switzerland*

<sup>2</sup>*Politecnico di Milano, Department of Electronics and Information (DEI)  
Piazza L. da Vinci 32, 20133 Milano, Italy*

<sup>3</sup>*Politecnico di Milano, MultiChancePoliTeam (MCPT)  
Piazza L. da Vinci 32, 20133 Milano, Italy*

## ABSTRACT

The paper presents a new multimodal 3D education environment for children with autism. The new multimodal interaction system considers a combination of visual, voice, and textual modalities. In particular, it allows children with autism to access contents through easy iconic symbols designed to guide them into the innovative environment. For that purpose, it has been very important to consider and identify the classes and attributes necessary to correctly describe different users. In the architecture hierarchy three different user profiles have been considered and structured, following the ICF\* model (an extension of the WHO International Classification of Functioning, Disability and Health guidelines), and describing both static and dynamic properties. A specific iconic language has been used to enrich and to present the virtual environment. Simultaneous visual, audio, and cognitive stimuli have been carefully used: they could be potential barriers but also rich opportunities for persons with autism. It has not been only a matter of putting information in a virtual space; it has been necessary to design and develop new languages, metaphors, and codes of interaction, in order to reduce the distance between the user and the system. In this case, communication talks via images, sounds, and gestures have been fundamental. The approach of the project takes into account the user model, the user profiles, the personalization, and the experimentation.

## KEYWORDS

Multimodality, Interaction, Accessibility, Autism, Learning, Virtual Reality

## 1. INTRODUCTION

Since the 21st century, social life has been conditioned by the introduction of new technologies. In particular, Virtual Reality and 3D collaborative systems permitted to create and compose powerful virtual communication environment. For that reason, it is interesting to note that when children with autism start using new technologies to communicate, they often share new and fundamental opportunities.

The paper is based on two existing projects. The first project is called WEBminore [9] (see Figure 1), and was developed by the University of Applied Sciences of Southern Switzerland ([www.webminore.supsi.ch](http://www.webminore.supsi.ch)); the second project, named ICF\* [26], has been developed by the Politecnico di Milano. The goal of the present work has been to define a new multimodal interaction interface for the WEBminore environment, leveraging the ICF\* model, in order to adapt the system to the needs of users with autism.



Figure 1. The WEBminore environment overview

Taking care of the multimodality of the realized environment and of the relational context, children and adults with autism can have the unique opportunity of bypassing some of the difficulties they encounter during their social life, such as visual contact, unwieldy movements, and face-to-face interactions. From this point of view, shared 3D virtual environments with multimodal interaction interfaces are an important opportunity able to improve the social life of persons with autism.

The main goal of this work has been the definition of a methodology of multimodal interaction that could allow users with different profiles of autism to play on an equal level with other users in a 3D virtual environment. In particular: a) we have identified specific techniques needed to help persons with autism to get involved and oriented in a 3D educational environment; b) we have used the innovative ICF\* model to profile different users; c) we have realized a prototype offering, to the different profiles of users, personalized scenarios; and d) we have involved some persons with autism to make a first validation of the methodology and of the realized environment.

In particular, focusing on the purpose of the project, the entire structure of the work has been characterized by the following pillars and innovative aspects:

- **User model**

The project leverages ICF\*, an extended version of the International Classification of Functioning, Disability and Health, with particular focus on attributes related with new technologies (mouse clicking, pointing, typing, etc.) and with different profiles of autism.

- **User profiles**

For the project we have considered three specific profiles of autism: 1) a profile with high thinking and recognizing signs problems; 2) a profile with voice production and verbal communication problems; and 3) a profile with social impairments and restricted patterns of behavior.

- **Personalization**

The project takes into account the personalization of contents (exercises and mini educational games), the input/output interaction, the navigation in the 3D environment, and the multimodality approach strictly related with the user profiles (text, voice, and video).

- **Experimentation/validation**

The prototype of the application has been tested with the collaboration of Esagramma ([www.esagramma.net](http://www.esagramma.net)), an Italian centre specialised in mental disorders and rehabilitation. The test involved six different users with six different levels of autistic disorder (belonging to the three defined profiles).

Autism classifies a set of complex neurodevelopment disorders, characterized by social impairments, communication difficulties, and restricted behaviour patterns. According to Simon Baron-Cohen, autism is the most weighty childhood psychiatric pathology in the world [5]. It is not a common pathology, however, as only four to fifteen children out of a thousand are affected. It exists all over the world and in all social classes but, as reported by the Diagnostic and Statistical Manual of Mental disorders [13], autism is a phenomenon that is four or five times more frequent in males.

## 2. AUTISM, INTENTION AND EMOTION

The difficulty to understand others' mind seems to be one of the big cognitive feature of autism and of a variety of autistic mental states such as: beliefs, desires, intentions, imagination thoughts, and emotions [5]. There are two main behaviours that are considered to be a forerunner of the theory of mind: shared attention and proto-declarative intentional communication.

Shared attention consists in children's behaviour around the age of nine months, when they start showing interest in all the objects seen by adults; they change continuously their view from one object to another and point to the adult beside them. The act of pointing, and the related natural gestures, characterize by definition the origin of the human communication. At this age children start building the common ground with artifacts present in the surrounding environment.

Proto-declarative intentional communication represents the same behaviour described above, but in this case the intention of the children is to communicate something; children point directly at objects and look back and forth between the objects and the adults, until the adults look at the same objects they are looking at.

In both cases, children are not only interested in influencing others' behaviour to obtain a concrete objective; they are also interested in changing the internal mental states of others. In fact, they aim to obtain and create a shared experience. Children with autism have deficiencies on both processes, because they are not able to develop a correct theory of mind.

The Baron-Cohen hypothesis (1985) states that children with autism do not develop capability to conceive the fact that other people can know, have desires, can hear and believe in something. This meta-representative deficit is the main reason for their problems related to social behaviour and verbal communication. For what concerns the emotional disorders, there are two different positions: the first one refers to a classical approach related to the theory of mind, which emphasizes the problem of understanding personal and mental states of other people [32]. The second position considers emotional disturbance as the primary disease for children with autism [16]; the origin of the problem is thought to be biological and children could be unable to perceive the emotions of others. As a consequence they do not know how to share the common environment and its social view [19].

### 2.1 Autism in Relation to New Technologies

It is interesting to note that when children with autism start using new technologies on the net, communicating by means of keyboard and mouse, they do not have the difficulties they encounter during their social life (visual contact, unwieldy movements, and face-to-face interactions that are not visible by others). From this point of view, the Internet could be an interesting new channel that could improve the social skills of people with autism.

For example, using virtual reality for children with verbal communication problems could be a great opportunity to aid those children to communicate with other users. These people could interact by using a keyboard and a mouse, and a specific iconic language that guides them into the system by means of a multimodal interaction environment. In particular, iconic communication is primary and fundamental; a set of iconic images used to access the system could possibly help people with autism to get information they require.

A dedicated multimodal interaction interface within the system could be created specifically for children with autism, encouraging them to interact with teachers and pairs for longer sessions of work. The multimodal interaction interface, based on an iconic language, could be used to improve the visual communication among users and between users and the system. Images could be used to help children with autism during the communication process, especially with memory, language, and speech deficiencies. This way, communication could not only be enhanced for people with cognitive degeneration or impairment, but also for people with normal cognitive levels.

Persons with autism have problems in understanding where their bodies finish and where the external world begins. A person with autism reported that he was not able to "find his body". Another person with autism described a fragmentary type of perception that she had about her body, whereby she was only able to perceive a single part of her body at a time; she had a similar fragmentation even when she tried to look at the environment surrounding her: she could only watch single object at a time.

Another important aspect related to the body is the “sense of touch”. For people with communication problems a heightened sense of touch could be extremely useful to learn things, independently from their form and content. For example, a person with autism said that he learned to read by touching the physical letters placed on the ground.

Finally, another interesting aspect related to the body takes into account the relief that people with autism feel when a part of their body is pressed. In particular, persons with autism have a strong desire to create pressure on their body in order to eliminate some kinds of pain.

### **3. THE 3D EDUCATIONAL ENVIRONMENT**

An interactive 3D environment allows individual communication in a shared space; users are involved in the same virtual space, but are physically separated. A collaborative and interactive 3D environment allows teachers and students to exchange learning opportunities in 3D virtual worlds. These multi dimensional virtual worlds are useful to provide powerful platforms and an ideal environment for teaching, learning, brainstorming, and sharing experiences. From an educational point of view these new technologies offer plenty of opportunities to create new information, to retrieve existing knowledge, and to build a common educational pattern.

The impact of virtual environments in this field creates many new challenges. One would no longer see pictures or frames posted on a flat wall or on a classical blackboard; instead, they will be published in a virtual space on which it would be possible to explore the contents in multi dimensional mode, and to navigate and search for information in an immersive system. This is a completely different approach, which offers the possibility to involve students in an environment and in an information context. Embodying them as a part of the system allows a set of functionalities that are useful in participating and contributing new experiences and information to the system. In essence, it is a new multidisciplinary technique that considers the fact that the role of students is also to contribute, promote, and sustain educational activities, sharing time in the same virtual space but allowing people to choose different behaviours. In this case the process needed to exchange information is not linear but reticular (information spatially connected in a 3D cyber world) and always evolving.

#### **3.1 Interactions, Presence and Perception**

Being in a virtual reality and interacting with objects and artifacts, users can experience the so-called “sense of presence”: something that establishes feelings and involvement around the user in a virtual space [17]. The development and the diffusion of computer interactive systems, such as 3D virtual reality, have been characterized by an increasing perception of the importance of the sense of presence. A technical definition of the sense of presence is “the subjective feeling of being there” [10].

The sense of presence is fostered by user’s ability to modify the environment according to her/his goals, and this is strictly related to actions that the user can perform in the system and the corresponding feedbacks it generates. Perceiving the effects of her/his actions, the user gets the sensation that the environment is considering her/his as “present” [18]. Moreover, the possibility that the user feels that the action is happening in “first-person” (and not shown in a third-person camera angle) is another crucial factor that determines presence.

The greater the user’s ability to reach a high level of presence within an activity, the better her/his involvement in the interaction will be [24]. For this reason, if the user is “present” in a space and is able to perform actions and perceive feedbacks, it is possible for her/him to successfully transform intentions into interactions and communication.

Developing 3D applications based on the interaction between objects needs attention; the focus should not only be on the quality of the graphics, but also on the cause-effect relationship, in order to support the opportunities of interaction (affordances).

## 3.2 User Centred Approach

Designing an artifact that aims to embody emotional experiences in interaction is extremely hard; yet this must be done when interacting with users [14]. For users to get in touch with their emotions and with what they want to express with their message, it is important that they are not forced to simplify their messages or their graphical expressivity. In this case the user centred approach is another key element used to design the system and its graphical user interface. This is directly related to the user experience: the way people feel about technology whereby look at it, hold it, and open or close it.

Visibility, accessibility, legibility, usability, language, feedback, and constraints are essential elements of the user centered design approach. In particular, visibility is an important characteristic needed to build the user's mental model of the system. As suggested by Kenneth Craik in 1943, models help users to predict the effects of their action while using the system.

Users should be able to understand immediately and intuitively what they can and cannot do within the system. For example, avoiding invisible and automatic system controls make it easier for all kinds of users to understand and interact with the system. Moreover creating an efficient multimodal interaction system, based on iconic languages, is useful to navigate the system's contents, and to improve the entire virtual interaction / communication.

Often, feedback is one aspect of the system that is undervalued by designers and developers. On the contrary, we argue that using specific sounds, labels, highlighting, animation, or a combination of any of those, could strategically enhance the system's affordance.

## 3.3 Related Works

Among other papers we analyzed and compared different innovative 3D systems; these systems helped us during the design and implantation phases of our work.

The "POWER UP" project [30] is a multi-player virtual world educational game, with a broad set of accessibility features. It is an interesting 3D virtual game –focused on perceptual, physical, and cognitive disabilities– which studies in deep the "sense of self" of the user.

The "pOwerball" computer game [8], another interesting project for children aged 8-14, is designed to bring together children, with and without physical or learning disabilities, encouraging social interactions surrounding the play. The approach used by this project was interesting for us because it put in the middle of the entire work the involvement of the children, in order to identify their behaviour during the game and its rules.

Finally, the "Minus Two" project [23] was very interesting in order to understand how an effective approach related with assistive technologies and interactive narration could be. As our project focused on multimodality, "Minus Two" was useful to learn how they managed to provide in the same environment the following features: listen, navigation, playing and orchestrating.

Other projects have been considered to help the design of our interactive environment. In particular, project [1] describes problems related to the navigation of virtual environment for people with motor disabilities; the project takes into account even problems belonging to the area of children with behaviour impairments. Project described in [20] takes into account the relation between the 3D user interface and video games, giving a set of examples that show the positive and negative aspects of both; the project was interesting for our work, especially for what concern the 3D special interaction and its complexity. In [18] authors treat the virtual reality as a training tool for disabled persons, where: a) a 3D environment can employ images rather than number and text; and b) the illusion of position, the illusion of depth, and the interaction with simulated environments can be used. The project described in [11] manages different kinds of information –voice, video, and data– and suggests a real-time visual-interaction for vision-disabled persons. Finally, in [21] authors discuss communication by means of a vocabulary of images and icons; it is the same strategy used in our work in the field of autism and its impairments.

## 4. THE NEW 3D SYSTEM ARCHITECTURE

As a 3D virtual environment that is accessible through the traditional input model (keyboard and mouse), the new multimodal interaction system has to consider a combination of visual, voice and textual modalities (speech recognition, speech synthesis, and recorded audio for output). The design process has to consider many different modalities to access the contents, and this multimodal user interface has implications on accessibility. For example, the user interface should allow users to access a textual content by pressing an icon and by choosing (with a mouse) a preferred output modality: audio, video, text or video. The system behind this should be able to distribute the selected content in the requested form, and in case the content is not available in that form, it should give back to the user a comprehensible feedback.

In particular, the multimodal interaction for this project combines the vocal modality and the regular, visual, point-and-click modality through mouse and keyboard. We defined all the elements necessary for managing the user profiles and building up a functional and accessible graphical user interface, with a multimodal interaction, for users with autism.

As a first step, we followed the ICF\* guidelines to design new classes needed to manage correctly the user profiles; secondly, we integrated them in the final and new architecture to be able to obtain personalized scenarios. This process is a sort of refactoring on existing functionalities without applying changes on what is already working. It is very important to identify all the classes and attributes necessary to manage correctly all kinds of user profiles.

### 4.1 User Model and User Profiles

In the architectural hierarchy each user is described from two different points of view: static and dynamic. The static description is used to manage those properties that are statically set by the user through a dedicated form or interface. The dynamic description is used to characterize data that is collected while using the application, and that depends on information describing applications and devices (see the MAIS Reference Model, described in [29]).

The five ICF's categories and their attributes are necessary to design the new architecture and to manage all possible user profiles without discrimination. For example, omitting one of these sections could generate some imperfection (lack) on the multimodal interaction within the system. In this way, it was possible to take into account features regarding memory, attention and emotional functions. But to take into account functionalities related to human-computer interaction, we adopted the ICF\* model and its ITF Skill attributes (typing, pointing, making gestures, voice recognition, and interacting with virtual and 3D environment). Figure 2 shows a graphical representation of the three user profiles defined and used to validate the project.

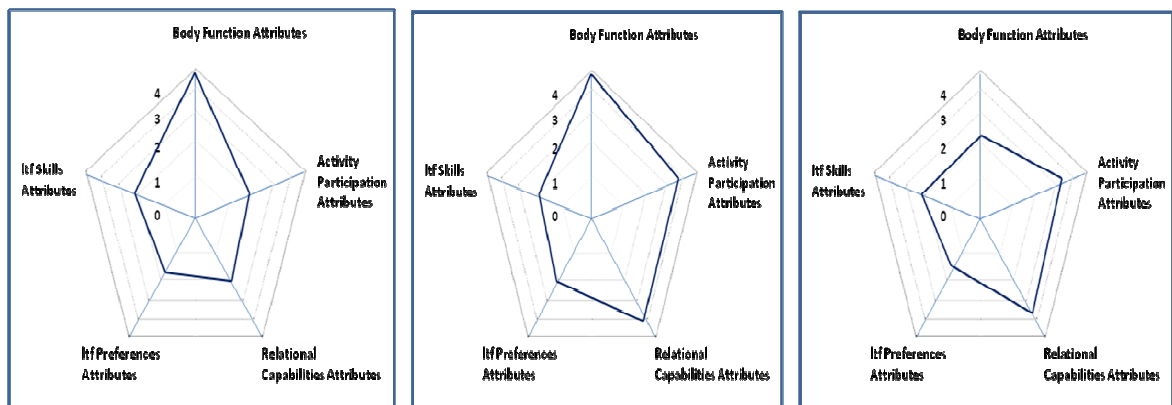


Figure 2. Three user profiles, with three different levels of autism

## 4.2 Personalization

The main topic has been to design and realize a prototype of multimodal interaction for users with autism in a 3D educational environment. The structure of the existing application has been modified to allow users with autism to access the application and to navigate in it with ease. Once the user finds a multimedia content, she/he has the ability to choose from a multimodal interface the kind of modality she/he prefers to use in order to access the content. Even though there are more modalities to access contents, the system is still able to adapt automatically the contents to the channel and modality chosen by the user profile. Three main modalities are considered: listening, visualizing and watching contents.

Users access the system by means of a dedicated graphical interface, on which the multimodal interactions are available according to the user model (see Figure 2). The main idea is to offer the user a dynamic multimodality interaction interface –based on simple, understandable metaphors– that is positioned around the multimedia content; such interface shows interaction modalities that are more suitable for the user. For example, Figure 3 shows the interaction interface associated to the “television” content.

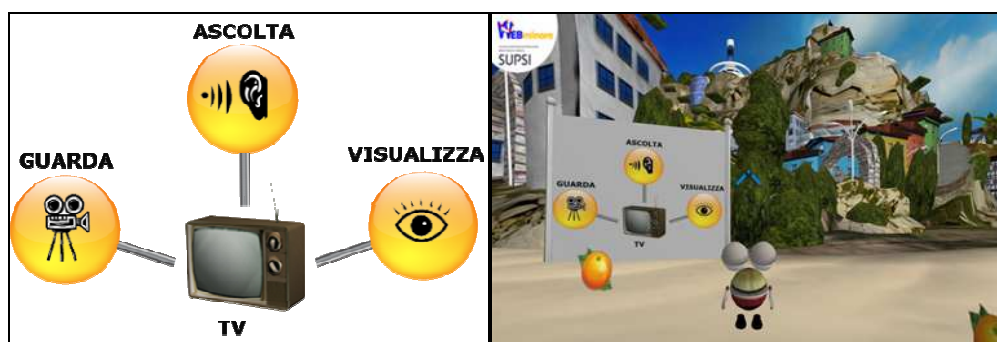


Figure 3. Multimodal interaction interface overview

A technique designed to guide the user in the system –without being lost or discouraged– takes into account the usage of FruitPath, a guided path based on “fruit” icons (see Figure 4, on the right). This kind of metaphor is useful to reduce the interpretation distance between the user and the system. Several FruitPaths are distributed in the 3D environment, in order to stimulate the user to move the avatar and discover the multimedia contents. We argue that giving the user an iconic path to follow could be a good strategy to eliminate all potential misunderstandings during the user’s interaction.

Another technique designed to help users during the navigation of the 3D system is based on WigWag signs: a set of specific signs, flags, and arrows useful to give information about directions. When users move into the system, they find such WigWags signs, which help them to discover the multimedia contents. It is very important to keep low the visual complexity of the signs used. Figure 4, on the left, shows some arrow signs we defined; notice that the design is quite simple, while the small fruit inside the arrows, and the colors used, recall the FruitPath.

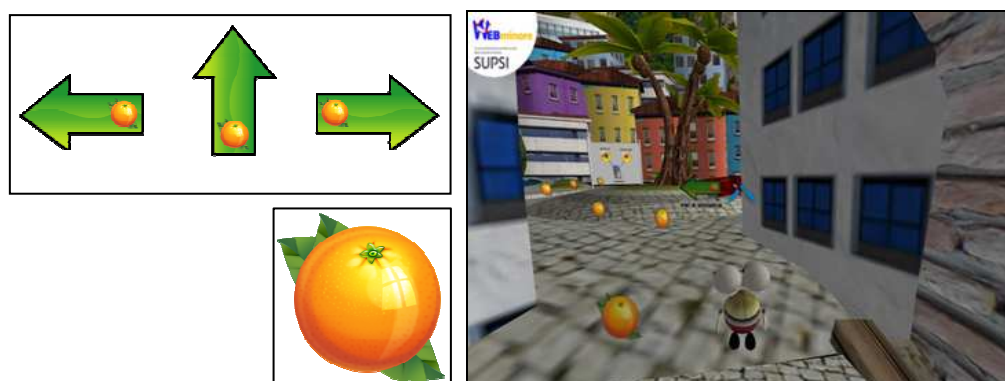


Figure 4. Example of the Fruits path metaphor, and a subset of the Wig-Wag signs used in the system



## 5. EXPERIMENTATION AND RESULTS

The prototype we built has been tested with the collaboration of Esagramma, an Italian centre specialised in mental disorders and rehabilitation. The test involved six different users with six different levels of autism (belonging to the three defined profiles). The approach used for the experimentation took into account a framework, based on user models, user profiles, and the ICF's guidelines.

The FruitPath was really appreciated by most of the users, and the creation of a simple and recognizable icon to follow in the system turned out to be a good strategic idea. The psychologist confirmed that all the involved users were explicitly attracted by the entire tridimensional system objects (environment), which gave them other interesting stimuli. The usage of the designed multimodal interaction interface was quite good and, after an explanation of the prototype and a short demo by the tutor, most of the users were able to recognize and use it by pointing and clicking their mice. Their perception and cognitive comprehension of what they were doing was not completely clear, but they did understand the simple multimodal icons shown by the system (listening, visualizing and watching). For what concerns the "WigWag" sings, not all the involved users were able to completely follow them in the system.

For time constraints only a few different contents have been produced, but very good results have been observed with respect to the personalization of multimodal existing contents, navigation, and interaction. Future work is going to realize more specific contents for specific multimodal channel to associate to the different user profiles.

## 6. CONCLUSION

One of the most important further remarks for this project takes into account the approach needed to use the 3D system with special devices, such as wireless remote controllers. It could be interesting to control the avatar without mouse and keyboard, but just a single device such as, for example, the Nintendo's Wii controller. This way, users could use a single device, capable of moving the avatar and interacting with artifacts in the system. It would be a huge simplification, requiring less cognitive efforts and giving the user more flexibility for what concerns movement, coordination, and interaction.

Another interesting solution is controller-free gaming, where a particular device permits users to interact with the 3D applications without any game controller; the Microsoft Kinect is one of those devices. For what concerns accessibility, such kind of controllers, which allow interaction with other environments and other people by means of simple natural gestures, could be a great solution. A virtual environment that reproduces (one-to-one) specific situations of real life could be created. Here, the user with autism could be immersed by means of her/his physical presence in the space, and by using her/his limbs and voice she/he could interact with the virtual environment and all the persons connected to it.

Considering the good results of the entire project (approaches, techniques, methodologies and experimentation), it could be interesting to exploit them in order to design and build a more sophisticated, dynamic customization mechanism for the system, using the five ICF\*'s main axes: Body Function, Activity and Participation, Relational Capabilities, ITF Skills and ITF Preferences.

Finally, considering relational problems of autistic children and good interaction results obtained during the experimental sessions, we are planning to allow different styles of interaction and cooperation between children, their teachers and pairs. We are confident that a 3D multimodal environment will be able to offer unique opportunities also from that point of view, and that an enrichment of the realized modular architecture will make that possible.

## REFERENCES

- [1] N. Alm, John L. Arnott, Iain R. Murray, Iain Buchanan, Virtual reality for putting people with disabilities in control, Department of Applied Computing, University of Dundee, UK.
- [2] T. Barbieri, A. Bianchi, F. Carella, M. Ferrà, and L. Sbattella, 2005, MultiAbile: a multimodal learning environment for the inclusion of impaired e-Learners using tactile feedbacks, voice, gesturing, and text simplification, Assistive Technology shaping the future, Lille, France.

- [3] T. Barbieri, A. Bianchi, L. Sbattella, 2004, "Minus Two": a cooperative and multimodal virtual environment to communicate with children with autism and mental diseases, Paris, France.
- [4] S. Baron Cohen, 2001, Theory of mind in normal development and autism, University of Cambridge.
- [5] S. Baron-Cohen, 1997, L'autismo e la lettura della mente, Astrolabio.
- [6] C. Bellettini, 2001, Theory of mind in normal development and autism, University of Cambridge.
- [7] C. Bellettini, A. Bianchi, L. Sbattella, July 2000, Distributed and Cooperative Story Telling for the treatment of cognitive and relational disorders, Int. Conf. on Computers Helping People with Special Needs, Karlsruhe, Germany.
- [8] B. Brederode et al., pOwerball: the design of a novel mixed-reality game for children with mixed abilities, Eindhoven University of Technology, Delft University of Technology.
- [9] M. Caranzano, A. Trivilini, M. Mainardi, Settembre 2008, WEBminore, Secure Use of Internet for kids, From Human's Behavioural Point of View, International Congress Hong-Kong (ISPCAN).
- [10] A. Carassa et al., A Situated Cognition Perspective on Presence, Università della Svizzera italiana
- [11] J.-M. Chung et al., Virtual Laboratory Education for Persons with vision Disabilities, Oklahoma State University.
- [12] R. Cooper, 2009, Using Information Technology to Assist People with Disabilities, Bioengineering, Rehabilitation Science & Technology University of Pittsburgh.
- [13] DMV-IV work group, 2010, Diagnostic and Statistical Manual of Mental Disorders, American Psychiatric Association.
- [14] M. Despontin, L. Sbattella, and R. Tedesco, November 2010, Natural language processing for storytelling and role playing: a training system based on the Propp model, Proceedings of Int. Conference on Education, Research and Innovation (iCERi), Madrid, Spain.
- [15] P. Dourish, 2001, Where the action is, The foundations of embodied interaction, MIT.
- [16] R. Hobson, 2006, Developing Self/Other Awareness: A reply, The Society for Research in Child Development, Inc.
- [17] M.P. Huang, N.E. Alessi, 1999, Presence as an emotional experience, University of Michigan.
- [18] A. Jacqueline, Chestnut and Lesia L. Crumpton, Virtual Reality: A Training Tool in the 21 Century for Disabled Personas and Medical Students, Mississippi State University.
- [19] R. Jordan, 2005, Chapter nine, Autistic Spectrum Disorders.
- [20] JJ. Laviola Jr., Bringing VR and Spatial 3D Interaction to the Masses through Video Games, University of Central Florida.
- [21] X. Ma, Jordan Boyd-Graber, Sonya Nikolova and Perry Cook, Speaking Through Pictures: Images vs. Icons, Princeton University.
- [22] G. Mantovani, G. Riva, 1999, Real presence: How different ontology generate different criteria for presence, telepresence, and virtual presence.
- [23] M. Principe, L. Sbattella, Minus Two – Un ambiente virtuale per lo storytelling musicale, Politecnico di Milano.
- [24] G. Riva, 2008, Enacting Interactivity: The role of presence, IOS Press
- [25] L. Sbattella, 2006. La mente orchestra. Elaborazione della Risonanza e Autismo. Vita e pensiero, Milano.
- [26] L. Sbattella, R. Tedesco, and G. Pegorari, March 2011, Personalizing and making accessible innovative academic services using ICF\*, an extended version of the WHO ICF model. Proceedings of the INTED conference. Valencia, Spain.
- [27] L. Sbattella, 2010, Accessible ICT for Vocational Education and Employment, 4th European eAccessibility Forum, Paris, France.
- [28] L. Sbattella, R. Tedesco, 2004, Profiling users in "Virtual Campus". The adaptability of an advanced and distributed learning environment, Istanbul, Turchia.
- [29] B. Pernici (ed.), 2006, Mobile information systems, Springer.
- [30] S. Trewin, L. Hanson, R. Laff, A. Cavender, POWER UP: An Accessible Virtual World, IBM Tech Center and University of Washington.
- [31] A. Whiten, 1993, Evolving a theory of mind: the nature of non-verbal mentalism in other primate.
- [32] Uta F., J. Morton, A. Leslie, 2003, The cognitive basis of a biological disorder: autism, MRC § Cognitive Development Unit London.

# GENERAL AND AUTOMATIC EFFICIENCY TESTING FRAMEWORK FOR JAVA

Narin Işık and Kerem Önal

*The Scientific and Technological Research Council of Turkey,  
Space Technologies Research Institute, Ankara, Turkey*

## ABSTRACT

When an application is claimed to be efficient, this assertion must be confirmed with valid metrics. Besides, these metrics are required during design iteration planning. User-based approach for usability evaluation may supply these necessary metrics, if its cost can be afforded. Its cost is high since it is done manually. This study sets forth a framework for efficiency evaluation that takes inspiration from the mentioned approach. The framework contributes to that approach in terms of efficiency by automating and generalizing it for any Java / Eclipse Rich Client Platform (RCP) application. The results presented here indicate that, the framework supplies valid efficiency statistics that can be used to plan redesigns and satisfy the efficiency criterion preset for the application.

## KEYWORDS

Human Computer Interaction (HCI) aspects, efficiency evaluation, usability measures, framework, Java.

## 1. INTRODUCTION

The four crucial aspects of HCI are (1) interdisciplinarity (it is related with computer science, psychology, sociology, education, etc.), (2) impact on human, (3) design and (4) usability (most central concern) (Shackel, 1997) as seen in **Figure 1**.

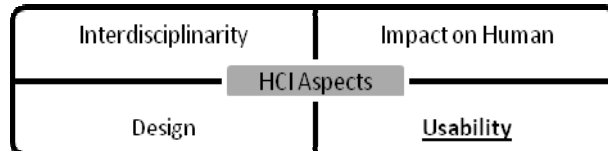


Figure 1. Crucial aspects of HCI

Usability has varying dimensions in different definitions, but efficiency is common in most definitions (Nielsen, 1993) (Shackel, 2009) (ISO, 1998). International Organization for Standardization (ISO) defines usability dimensions as (1) effectiveness, (2) efficiency, and (3) satisfaction (ISO, 1998) as in Figure 2. This study focuses on how efficiency can be reported quantitatively whilst design iterations are ongoing.



Figure 2. Usability measures

When usability is evaluated, the approach (how data will be obtained) and type (purpose) should be identified (Nielsen & Mack, 1994) as in Figure 3. Applying user-based approach for all aspects/features of all design iterations is expensive. However, it provides the realistic metrics for efficiency. It is preferably realized

at the end of all design iterations due to its cost. On the other side, heuristic evaluation is cheap and rapid, but it is suitable for problem diagnosis. It is not enough when some metrics are needed to measure efficiency (Sherry & Macredie, 2005). Moreover, cognitive walkthroughs simulate user's problem-solving process (Nielsen & Mack, 1994). It never replaces the user based approach, since the metrics provided are predictions, not real performance eventually.

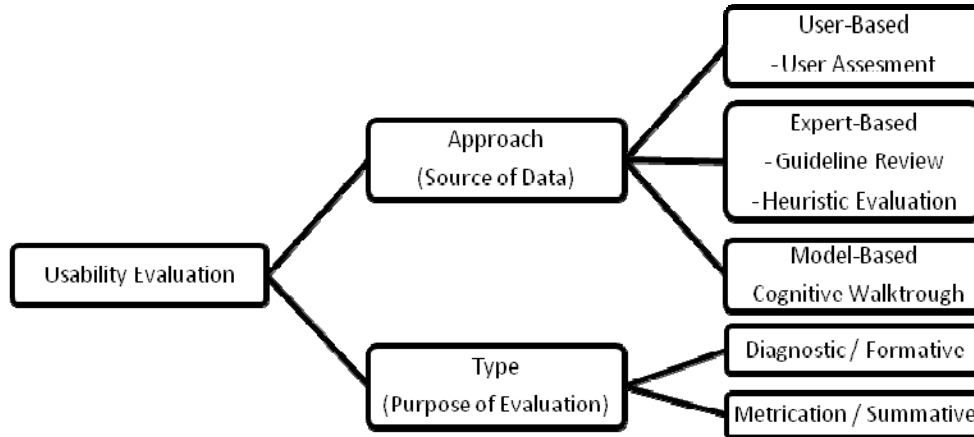


Figure 3. Usability-evaluation framework

The rationale behind this study is to facilitate the efficiency measurement of different features of each design-iteration by automating it. The manual tests such as videotaping the users' test sessions and analyzing the content; or observing the user and using a chronometer to measure performance of tasks are more expensive. The aim of the study is to relieve the tester of these manual tests, which consume time and necessitate human resource, equipped lab environment, etc. The main consideration of the study is that, the efficiency of an application can be concluded from the logs generated by the application.

In this study, a general and automatic efficiency testing framework for Java is proposed. This framework is basically capable of recording all types of graphical user interface (GUI) events and concluding the efficiency of the tasks from these records. It is general, since only one line of code is added in plain Java applications and none for Eclipse RCP applications (McAffer et al., 2010). This keeps the application code clean. Also, the framework eliminates the need for logging additional tasks for new tests at redesigns. It prevents test repetitions when tester decides to change (split, merge, etc.) task definitions. Hence, the tester does not wish that the developer had logged more. Additionally, the framework is automatic, since it records user actions in a predefined format, separately from and independently of the application; evaluates efficiency of tasks; and provides cumulative results by its own. The framework captures exclusively GUI events since any other event can cause to dependency between the framework and the application.

## 2. METHODOLOGY DESIGN

### 2.1 Conceptual Framework

A framework named EfficiencyAnalyzer is developed which acts as a plug-in for Eclipse RCP applications and as a sub-project for Java applications. As depicted in Figure 4, the framework monitors user actions at the application it is added/plugged in. Task definitions can be defined and updated as a patch (McAffer et al., 2010). All user actions are recorded throughout the/a test session by the framework. Before the application ends, the framework searches for tasks according to predefined task patterns throughout the recorded user actions and evaluates efficiencies of the found tasks. Finally, efficiency statistics for tasks are generated.

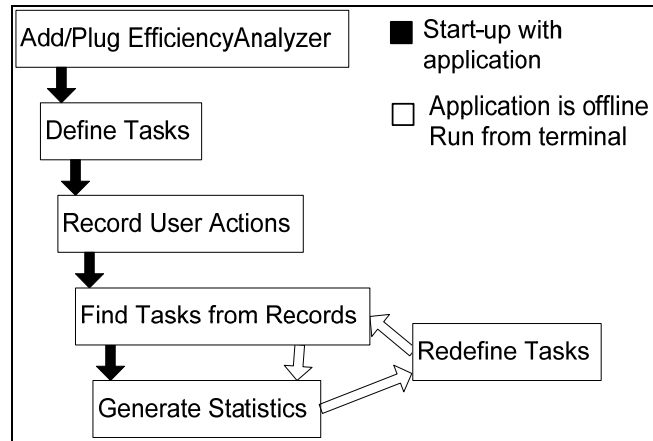


Figure 4. Conceptual framework

Additionally, the framework can generate statistics offline (when application is not running) to create a new report when the developer decides to redefine the tasks. In that case, test is not repeated since user actions were once recorded. This brings a huge benefit in terms of time and cost during application development. Only the statistics report is updated according to the newly defined tasks.

## 2.2 Logical Structure

As users start to use the application and perform some tasks, user actions such as mouse click, mouse move, key stroke, etc. are reflected as GUI events in the application. A GUI widget has listeners for different event types in Java. Both, Standard Widget Toolkit (SWT) and Abstract Windows Toolkit (AWT) / Swing Application Programming Interfaces (API) provide general mechanisms to listen all occurrences of an event type that is triggered by any GUI widget.

The framework, proposed in this study, has *a recorder component*, which records all predefined types of AWT, Swing, SWT and Eclipse RCP events of the application it is incorporated in (Figure 5). GUI event types, which will be recorded by the framework, such as AWT.MOUSE\_EVENT\_MASK, SWT.MouseDown, SWT.Close, etc are defined in a configurable file. One should pay attention to the fact that recording events such as mouse motion, mouse-enter, mouse exit may degrade performance.

Records are kept in Comma-Separated Values (CSV) formatted log files. A record contains information about event type (Widget event/Workbench event [page open, close, activate, etc.]), date, source type (Label, Text, Button, ViewSite, etc.), widget value, page name, window name, coordinates of event, etc. Optionally (not necessarily), the tester can analyze these logs using spreadsheet programs and running macros on them.

Additionally, the framework has *a reporter component*, which searches for tasks in log files and reports them in a CSV formatted file (Figure 5). Tasks are defined as GUI event patterns that consist of a start and an end event. For example, suppose that there is a button on page A. The user clicks on the button and the application runs a job (database operation, image processing, etc.). If the job is successful the application passes to page B; otherwise it stays at page A and displays a warning dialog. In order to evaluate the efficiency of the job when it is successful and unsuccessful the following tasks can be defined:

Table 1. Task definition examples

Task Name	Start Event	End Event
Task1	clicking the button	opening page B
Task2	clicking the button	displaying warning dialog

Generated report provides the following statistical data for each task:

- Average time taken to complete defined task
- Frequency of the task
- Average time of recurrence of the task

- Histogram of task durations. These data help to see whether the average duration of a task is in the time interval of the most frequently seen task duration. For example, average duration of a task can be 25 sec while most of the occurrences of that task are performed in 15-18 sec time interval. Hence examining time duration histogram diagram provides more insight about the efficiency.

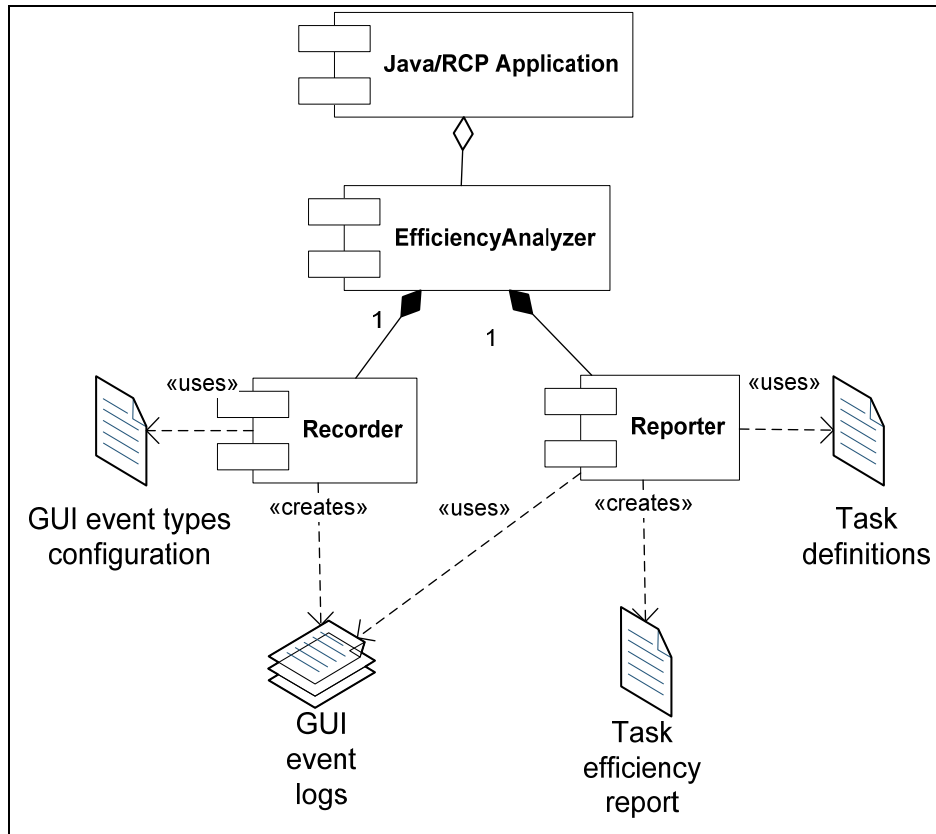


Figure 5. Component diagram of EfficiencyAnalyzer

### 2.3 Projection to HCI System Components

An HCI system has four principal components: user, task, tool, and context (Nielsen, 1993) (Shackel, 2009). During the development of EfficiencyAnalyzer these components were considered as listed below and are depicted at Figure 6:

1. *User*: The definition of efficiency evaluation (Nielsen, 1993) necessitates users to be expert. Before starting tests, it was assured that the users were experts. They experienced the application by completing 260 tasks to attain expertise level. During tests, it was not needed to interfere with the users by the help of the EfficiencyAnalyzer. Hence, reliable and convincing results were obtained at the end of the tests.

2. *Tool*: At first, the purpose of this study was only to measure efficiency for a specific application. Then, it was realized that the process was common for many applications: user performs tasks, GUI events are triggered, events should be logged, and tasks can be tracked from logs and reported. That provided motivation to this study to generalize the process as a framework and use the framework for different tools.

3. *Task*: Since tasks are specific for tools and users, they were defined as GUI event patterns in a file as tools changes. It was also made possible to redefine them, when tester needed to check additional tasks without repeating the test session with the users.

4. *Context*: One of the benefits of this framework is that it depends only on log files. This makes it independent from the context in which the user performs tasks using the tool.

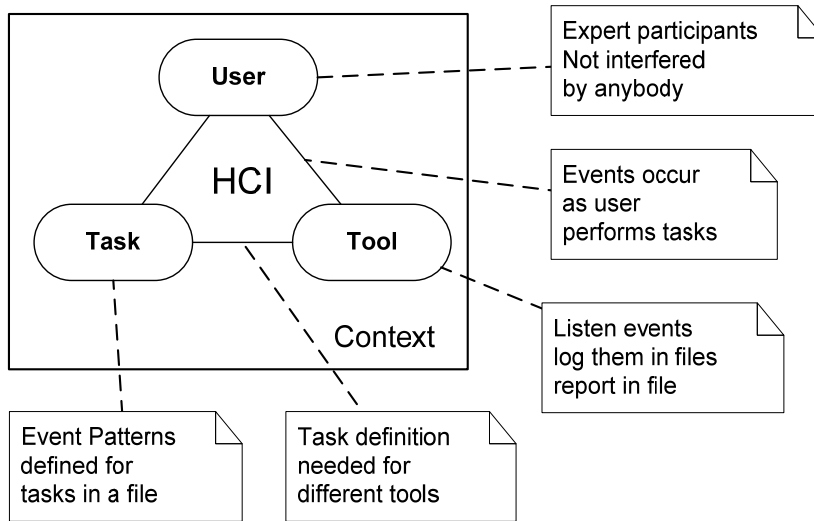


Figure 6. HCI system components

## 2.4 Experiment

This framework was used to conduct efficiency tests for an Eclipse RCP application. Tests were conducted after scheduled milestones (i.e. design iterations). Users were requested to perform specific number of jobs using the application. 6 tasks, that were involved in that job, were defined (named as *Task1*, ..., *Task6*). During a test, user actions were recorded by the framework. After a test session, all tasks were evaluated and a report was generated for that milestone automatically. Then, the report that summarized the efficiency of these tasks was analyzed by testers and developers.

For example, the report provided statistics for *Task1* as below:

- Average time to complete the task: 21 sec
- Frequency of the task: 45 occurrences in 1 day
- Average time of recurrence of the task: 25 min 2 sec
- From the histogram, it was seen that the task was performed mostly in 15-21 sec. It was in agreement with the calculated average time (Figure 7).

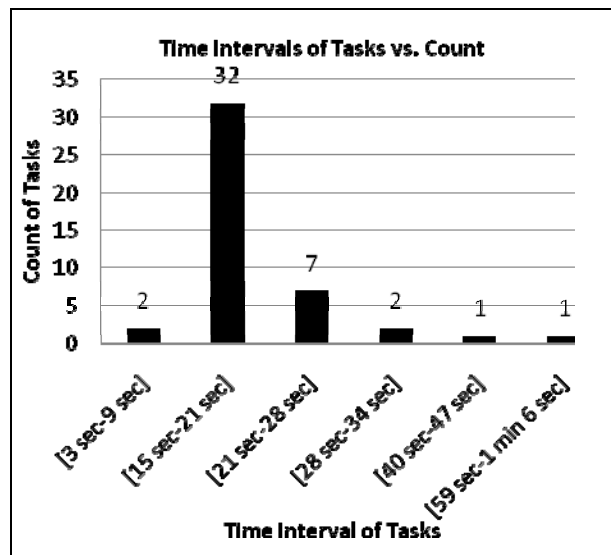


Figure 7. Average duration is most frequently seen duration

Statistics for *Task2* were as below:

- Average time to complete the task: 22 sec
- Frequency of the task: 41 occurrences in 1 day
- Average time of reoccurrence of the task: 26 min 51 sec
- From the histogram, it was seen that the task was performed mostly in 15-19 sec. It was less than the calculated average time (Figure 8). It was concluded that, it rarely took more time than expected. The users sometimes performed that task slower than they usually did.

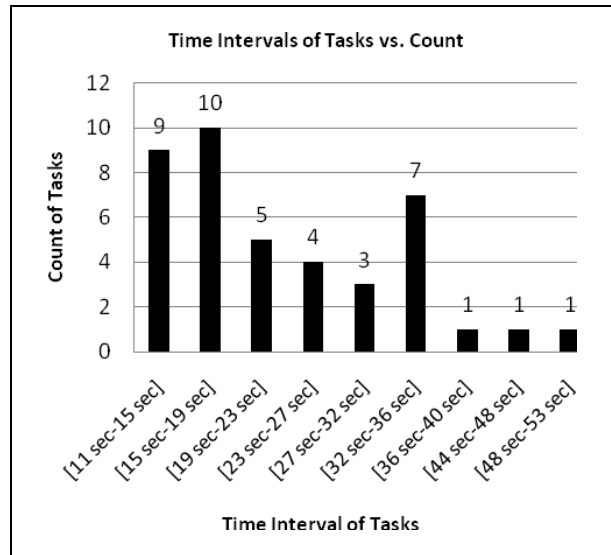


Figure 8. The average duration is not the most frequently seen duration

At the end of the all design iterations, statistics for all milestones was gathered as illustrated in Figure 9.

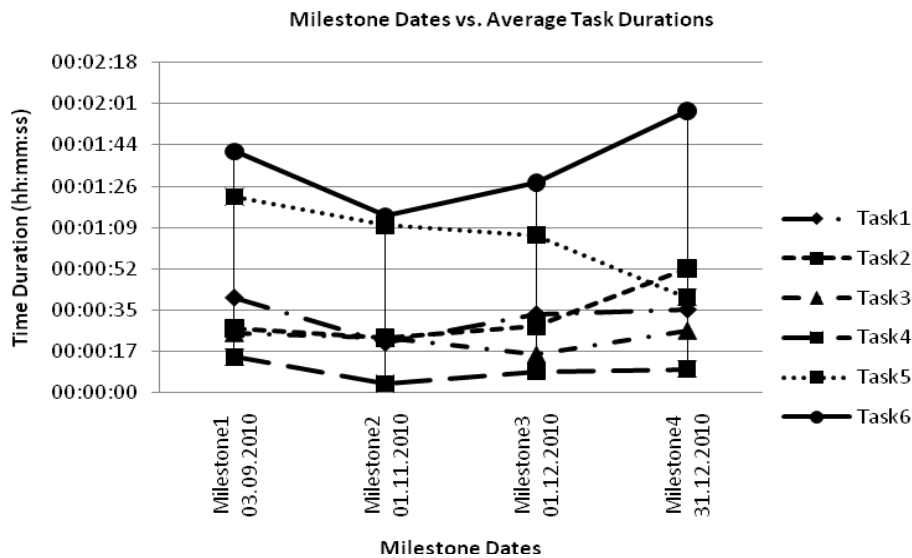


Figure 9. Final results for design iteration

It was observed that the costs of *Task1* and *Task2* were increasing while the cost of *Task5* was decreasing by design changes. Other tasks had slight changes in efficiency.

It is common that, the users of the application, request some efficiency limits before the application. The results obtained in this study also helped to check whether the efficiencies of the tasks were in acceptable



limits. The fact that, efficiency criterion for the application was met, was proved by the obtained empirical results.

### 3. CONCLUSION

The usage of EfficiencyAnalyzer decreases the number of efficiency tests dramatically as it eliminates the need of repeating tests when task definitions change for the same design. As all GUI events during the test session are logged, the tester can redefine tasks after the test and generate new measures from the same user action logs. Also, the framework eliminates the need of additional logging for different designs as it logs all GUI events.

Human resource allocation is not needed for efficiency measurement during the test sessions. Also, interfering with the users as they perform the tasks can be avoided. This leads to more realistic test results.

Efficiency tests were conducted to try to understand how the application could be made more efficient and how its usability could be increased. The results of tests were discussed among designers. These quantitative results helped designers to reveal potential problems, plan iterations for design and compare efficiencies of different versions of the product. Hence the ultimate design was achieved as expected in usability testing (Dumas & Redish, 1999).

In our case; during these discussions, decisions for design changes were made based on numerical, empirical and reliable results. The time the users needed to perform some specific tasks were figured out. In that vein, it was decided to design some tasks to work in parallel at the background. If EfficiencyAnalyzer results were not available, design enhancements would not be dared. If efficiency evaluations were not done automatically by the help of EfficiencyAnalyzer, the cost of manual user-testing for design iterations could not have been afforded.

In the future, this framework can also be used to understand when the users reach to expertise level. Since results of expert users reflect the real usage, determining the level of expertise is important. The learning curve of the users (Nielsen, 1993) can be derived by enhancing the EfficiencyAnalyzer so that it summarizes the efficiency of tasks periodically. Hence, no guessing on the experience level would be needed to accept a user as an expert.

### REFERENCES

- Dumas, J.S. and Redish, J., 1999. *A Practical Guide to Usability Testing*. Cromwell Press, Great Britain.
- ISO, 1998. ISO 9241-11 *Ergonomic requirements for office work with visual display terminals (VDTs) - Part 11: Guidance on usability*. International Standard.
- McAffer, J. et al, 2010. *Eclipse Rich Client Platform*. Addison-Wesley, USA.
- Nielsen, J., 1993. *Usability Engineering*. San Diego, Academic Press, USA.
- Nielsen, J. and Mack, R.L., 1994. *Usability inspection methods*. New York, John Wiley, USA.
- Shackel, B., 1997. Human-Computer Interaction - Whence and Whither? *Journal of the American Society for Information Science*, Vol. 48, No. 11, pp 970-986.
- Shackel, B., 2009. Usability - Context, framework, definition, design and evaluation. *Interacting with Computers*, Vol. 21, No. 5-6, pp 339-346.
- Sherry, Y.C. and Macredie, R.D., 2005. The assessment of usability of electronic shopping: A heuristic evaluation. *International Journal of Information Management*, Vol. 25, No. 6, pp 516-532.

# INTERACTION ERROR BASED VIEWPOINT ESTIMATION FOR CONTINUOUS PARALLAX ERROR CORRECTION ON INTERACTIVE SCREENS

Bastian Migge<sup>1</sup>, Andreas Kunz<sup>1</sup> and Tim Schmidt<sup>2</sup>

<sup>1</sup>*Innovation Center Virtual Reality, ETH Zurich, Tannenstrasse 3, 8092 Zurich, Switzerland*

<sup>2</sup>*Palo Alto Research Center, 23333 Coyote Hill Rd Palo Alto, CA-94304, USA*

## ABSTRACT

Many interactive screens suffer from the incoherence between image plane and interaction plane. The resulting gap causes parallax errors that hinder a precise interaction with the system. For many reasons, this gap cannot be physically reduced any further, while a software correction is still missing. Thus, this paper introduces an observation model for a continuous automatic recalibration controller of the touch sensitive surface.

First, we show that the overall interaction error stems partly from the parallax error, which depends on the changing viewpoint of the user. Hence, a static calibration cannot overcome this error. Being not directly measurable, a continuously adapting correction controller sets the appropriate correction parameter based on updating the estimate of the user's viewpoint from the history of his interaction errors. To estimate the user's viewpoint in front of the screen based on the interaction error on the screen, we secondly investigate the correlation of the two domains from data of a user study, working on a large interactive screen with a significant gap between image plane and interaction plane. The correlation analysis shows significant differences in the interaction error stemming for different viewpoints, which allows the controller to infer the viewpoint. Finally, we model the results as a discrete observation model for the Partially Observable Markov Decision Processes (POMDP) correction controller.

## KEYWORDS

User Modeling, Continuous Calibration, Parallax Distortion, Recursive Estimator

## 1. INTRODUCTION

Interactive screens as man-machine interfaces [1] are becoming increasingly popular [SMART® boards, Wacom® screens, Microsoft® Surface®, Apple Inc.® iPad®]. Large electronic whiteboards [2], mobile phones and Tablet PCs (e.g. [3]) enable highly intuitive operation. The direct manipulation of the screen's content with the user's finger or with tangible user interfaces [15,33] needs a good coherence between the input position and the digital content. To enable a precise interaction on such interactive screens, they need a good spatial and temporal alignment between the displayed objects and the input tracking system.

Common calibration techniques focus on eliminating static geometric distortions by a keystone correction. Being deduced from an initial calibration process, the calibration parameters correct the interaction position statically to the corresponding display coordinate. However, a spatial distance between the screen's surface and the glass pane of the tracking overlay still leads to a parallax error (Figure 1). This causes the interaction position for a specific target being dependent on the user's viewpoint. Since the user moves in front of the screen when interacting with the system, the calibration has to be continuously adapted.

A distance of centimeters between image plane and interaction plane is common for security reasons on ATMs and kiosk applications [5]. Such stability reasons to withstand the mechanical loads from touch interactions cause an offset, which creates an optical distortion in terms of a pointing error that significantly impairs the user interaction [6]. Usually, this error is avoided by enlarging the targets on the screen. This approach limits the number of targets being simultaneously shown on the screen, and impacts the usability of the application due to increasing number of menu layers and screen pages.

The parallax error is influenced by user characteristics like height, viewpoint, motion, and arm length. Thus, a static calibration depends on the specific user and results in distortions for other users during a team

meeting for example. Even for the same user, the static calibration does not take into account his positional change over time, impairing the user's pointing accuracy especially on large whiteboards (see Figure 1).

Initiating a static recalibration before each user interaction can naively solve these problems. Since this is not a practicable solution, interactive whiteboard systems are only calibrated once at the beginning of the session. Here, a hidden and continuous recalibration while using the interactive system is required. Since the parallax error depends on the viewpoint of the user, which is not directly measurable, a model based correction controller uses the history of interaction errors to estimate the user's position.

Our work follows the approach in [9], which defines the basic model structure of a POMDP (Partially Observable Markov Decision Process) controller. It estimates the user's viewpoint by continuously receiving pointing errors for specific targets on the screen, and predicts the future position based on a model of the user's behavior to set up the correction parameter for the next interaction. Hence, the controller consists of a correction policy that defines the correction parameter for the estimates of the viewpoint, and three models: The process model [5] describes the user's behavior in front of the screen to predict his position; the reward model describes the effect of correction actions to the user; and the estimation model describes the correlation between interaction error on the digital surface and the actual viewpoint of the user.

Within this paper, we present the correlation between interaction error on the digital surface and the user's actual viewpoint, which we derive from a user study.

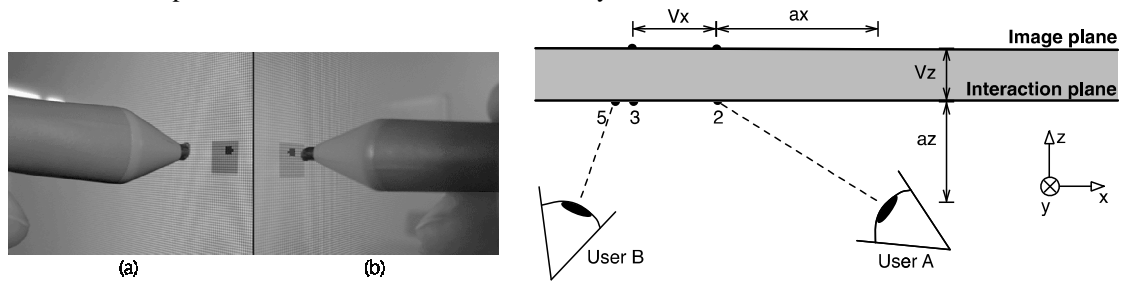


Figure 1. (a) 15 mm offset, (b) 7 mm offset between the LCD's image plane and the interactive surface of the input device. (In both cases, the tip of the pen contacts the interactive surface (a 6 mm glass pane), while the dot in the gray rectangle appears in the image plane underneath.), (c) Effect of the parallax distortion on the interaction point in x-direction.

## 1.1 Parallax Error

The parallax error shown in Figures 1 is caused by the offset between display- and touch plane. It affects any interaction on touch sensitive surfaces due to the close spatial coupling of input and output device. The touch plane is necessary to accomplish physical robustness and security for publicly deployed systems (e.g. ticketing machines). The likelihood of missing a target increases with the screen's size and the distance between image and interaction plane  $V_z$  (Figure 1(c)). Thus, the distortion depends on the distance between viewing point and the point of interaction, as well as on the viewing angle against the perpendicular.

It is:

- $V_z$ : Distance (offset) between image plane and interaction plane
- $V_x$ : Resulting parallax distortion in x-direction
- $a_x$ : Distance from user to interaction point in x-direction
- $a_z$ : Distance from user to interaction point in z-direction

Point 2 (Figure 1(c)) represents the interaction point of User A. Although he aims at point 4, point 1 will be registered resulting from the parallax distortion, mapped perpendicular onto the image plane. Point 5 shows the different interaction point for the same target, caused by the different viewpoint of User B. To avoid the parallax error, the user would have to interact at point 3. To reduce the effect of missing a target, the target size can be increased [21]. As a consequence, today's ticketing machines use large target areas to reduce the resulting interaction error. Due to the limited size of the displays, the number of targets is limited. This causes applications with cascaded submenu windows, which irritate the user and worsens usability [29][30].

Correcting the parallax distortion with a software controller can be easily done, once the user's viewpoint is known. The equation is:

$$V_x = \frac{V_z \cdot a_x}{a_z} \text{ (for the x-direction)}$$

Note that this equation only takes into account the geometric conditions and omits refraction indexes caused by the different optical media.

Since the offset can only be physically reduced to a certain account, an additional software correction is required in order for further increase the user's interaction accuracy.

## 1.2 Partially Observable Markov Decision Process

Enhancing the accuracy on interactive surfaces improves the precision of interactions, such as editing drawings and working with a widget-based user interface. A user-adaptive online re-calibration controller can overcome the parallax distortion, which depends on the user's viewpoint. Instead of tracking the viewpoint directly with high effort, the software component will use the history of interactions on the screen as evidence for the viewpoint to continuously re-calibrate the user-interface online. This problem can be formulated as Partially Observable Markov Decision Process (POMDP).

We focus on touch events from interaction on (or nearby) GUI elements that have well-defined target-points (e.g. small buttons). Comparing the offsets (from the center of these widgets) of several such interactions with a reference model of offset distributions for different viewpoints (finding of this paper) allows calculating a probability distribution over the different possible user positions relative to the screen. Conceptually, each of the possible positions represents a hypothesis. Once the software component has enough evidence supporting one of these hypotheses, it re-calibrates the interface for the corresponding viewpoint. We are hence trying to solve a sequential decision making problem.

Sequential decision making is a process (of potentially infinite steps) in which an agent at each step receives some information about the world and selects an action based on the accumulation of this information. He has to choose from a finite set of actions on how to interact with the world. In general, the information received is incomplete (hence does not allow deducting directly the actual state of the world), and the results of interactions can be uncertain. An optimal policy for such a process is a mapping of the accumulated information to action choices that maximizes the expected value of some valuation-function. A common formalization for such problems is a Partially Observable Markov Decision Process (POMDP in the following). Markov decision process models have proven to be useful in a variety of sequential planning applications where it is crucial to account for uncertainty in the process [7].

POMDPs can be described as a six tuple of States [8], a finite set of actions the agent can execute, and transition probabilities for successor states conditioned on predecessor belief state and action. The system is in one state at any point in time, which is not directly observable by the agent. This is modeled by a finite set of observations with occurrence probabilities conditioned by agent actions and system states. A reward function maps action-state pairs to values and a discount factor relates these values over different time steps.

The planner's goal is to select the optimal correction action based on previously accumulated information about the system's state. This accumulated information is represented as a probability distribution over all possible system states (referred to as the belief state of the planner/agent). Representing all information available to the agent at a point in time, a belief state fulfills the Markov Property [7]. An action choice is optimal if it results in a maximal expected value over all possible futures of the process.

To model the parallax-correcting problem as a POMDP, the 6 tuple is defined as follows [9]: The system's actions are the application of different adjustments to the parallax correction (or not). Its observations are the user interactions measured by the hardware's sensors for specific GUI elements, while (domain) states ( $S$ ) represent the different possible distortion errors caused by the user's viewpoint, which are not directly observable. The transition model describes the resulting distortion error after applying a correction action with respect to the current distortion error. The observation model probabilistically relates interaction observations to distortion errors. Finally, the reward model defines rewards for adjustments depending on the actual distortion.

## 2. RELATED WORK

Common calibration mechanisms for displays were initially driven to establish the geometrical alignment between display and tracking system. An approach of how to adapt the projector calibration to the projection surface is shown in [10] and [11]. However, these systems do not consider the influence of the user's changing viewpoint on the parallax error.

Today, interactive systems use flat screens instead of projectors. Although the compact design reduces the geometric distortions, the tracking system and the display system still have to be aligned. So far, the calibration is done in a dedicated process before using the system, in which the user has to hit a certain set of targets. The measured pointing errors are transferred into a static correction matrix, which is used to transform touch-coordinates to system coordinates [12]. Obviously, the parallax error cannot be corrected with such a fixed calibration if the user moves in front of the screen while using the system.

Reducing the offset between interaction plane and display plane is an effective approach to reduce the pointing error. As such reductions can only be reached by modifying the underlying hardware, they are naturally limited by the employed technology. A protective glass for example is necessary as it provides physical robustness of the interactive system, which is important for large interactive screens in public areas.

In order to further decrease the parallax error, only software solutions can be applied. Kent et al. [13] introduced a non-linear correction, which takes into account the tracking system's noise while Hong-Ping et al. [14] presented a learning calibration method based on back-propagation neuronal networks. Migge et al. [9] presented a general framework for a model predictive controller that increases the interaction accuracy by a user-adaptive online recalibration. Although they propose the problem as POMDP in general, the paper lacks of detailed analysis of the model parameterization. In [5], the process model is deduced from an empirical study, which models the user's movement in front of the screen.

Further studies investigated the effect of the target size on the interaction quality of touch sensitive surfaces. The ISO standard 9241-9 recommends the target size to be at least equal to the width of the index finger [19]. Jin et al. [28] investigated the target size and spacing on interactive surfaces for elderly. Other studies found that a larger target sizes led to better performance on a touch based numeric keypad [21, 22, 23, 25, 26, 27]. However, the literature lacks of analyzing the correlation between the interaction error and the user's viewpoint in front of the screen, enabling to estimate the viewpoint and correct the parallax error.

The hand-eye coordination can be evaluated with the Pegboard Test [20]. It is designed to motor dexterity and coordination of assembly line workers. Since the test is realized on a hardware board, it does not cover the task of interacting on an interactive screen. Fitts' Law [18] measures the difficulty of a given interaction task. Since the tests were designed for small displays, and due to the fact that we do not want a specific task to influence the user's interaction behavior, we do not apply this method.

## 3. CONTRIBUTION

In this section, we introduce the observation model for a POMDP controller to improve the pointing accuracy on interactive screens by correcting the parallax error. As we will show, the error stems from the viewpoint, which is not directly measurable. This interaction error (observation) is used to estimate the viewpoint (system state).

A common widget-based graphical user interface (GUI) provides a set of elements (e.g. buttons), whose events indicate the viewpoint. These so-called targets provide a reference point to measure (observe) the pointing error as the distance between target and interaction (touch) point. As Figure 2 (c) shows, the interaction error in principle depends geometrically on the viewpoint. Since the ability to hit small areas on the screen also depends on the user's dexterity, the measurements contain uncertainty. This uncertainty is described in the observation model. It expresses the correlation between the actual viewpoint and the resulting pointing accuracy (given as observation) as a discrete conditional probability distribution. It models the probability of an occurring pointing error (observation) for each viewpoint position relative to the target. Instead of observing the viewpoint directly, the controller uses the history of the resulting pointing error to estimate the viewpoint.

In order to set up the model realistically, we ran a user study monitoring the user interacting on a large screen with a significant offset between interaction and image plane, causing a parallax distortion based

interaction error. We measured the viewpoint of the user, the target, and the interaction position as three-dimensional (3D) data and analyzed the interaction error and the viewpoint relative to the target position. Based on the viewpoint relative to the target and the target's position on the screen, the correction controller can estimate the user's viewpoint in front of the screen.

We measure the interaction error of the test setup with its offset between interaction and display plane. Next, we show that the interaction error is partly biased by the parallax error: We compare the absolute pointing error, given by the offset between target and actual interaction position, with the offset from the assumed interaction position, given on the straight line between target and viewpoint. In preparation for the discrete observation model, we discretize the viewpoint space into 5 segments representing discrete system states. For each state, we show the discriminability of the observations, which is necessary to vice versa deduce the viewpoint from the interaction. Finally, we discretize the interaction space into a set of observations, and express the observation model as discrete probability distributions over the observations for each state.

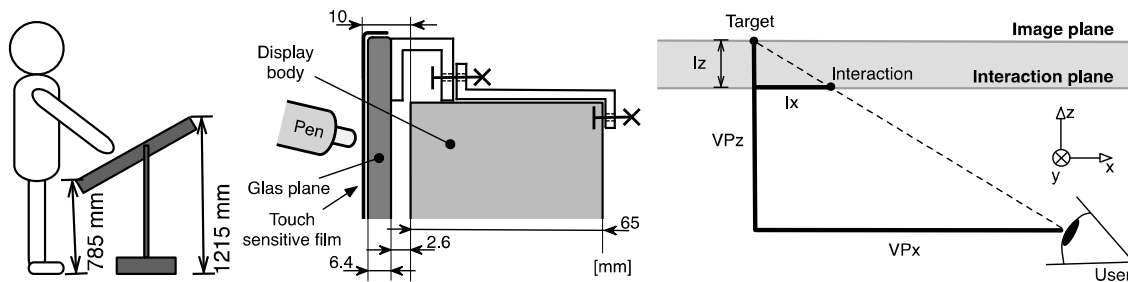


Figure 2. (a) Schematic measurement setup with test person (b) Assembly of the tracking system on top of the display. (c) Linear dependency between relative viewpoint (VPx) and interaction error (Ix).

### 3.1 Experiment Design

In order to model the user's interaction behavior realistically, the probabilistic model is deduced from empirical data collected in a user study. We measured the touch position on the interaction plane, the target position on the image plane, and the user's viewpoint position in front of the screen as 3D data, while the user executed a simple interaction task on a large interactive system.

#### 3.1.1 Participants

The average age of the 13 male and 4 female participants was 31.05, their median age was 29, and the average (median) height was 1790 (1840) mm. Four test persons were left-handed and 8 wore glasses.

#### 3.1.2 Task

As shown in Figure 2 (a), the user's task is to stand in front of the digital whiteboard, trying to hit a button that appears at random positions on the screen. Due to [19] we defined a realistic but small and symmetric target of 13 x 12 mm (15 x 15 pixels) to motivate the user to hit the target's center (see Figure 3 (c)). After each interaction, the button moves to the next position, regardless of whether the user hits the target or not. The user is encouraged to move freely in front of the screen during the task. The typical task duration (measurement acquisition time) is 5 minutes to avoid fatigue. The user's viewpoint is not only biased by his characteristics (height and arm length), it also depends on the position of the button element on the screen.

#### 3.1.3 Interactive System

We used a 50-inch plasma display (Pioneer PDP-502 MXE) with a resistive touch-sensitive overlay (SMART Technologies) mounted on top of the display. The display's resolution of 1280 x 768 pixels and its effective size of 1098.2 x 620.5 mm result in a pixel pitch of 0.858 x 0.808 mm. Each pixel is implemented as a square arranged set of cells for each primary color, with a distance of less than 0.2 mm. The front glass is divided into a front glass substrate, a dielectric layer, and a protective layer to stabilize the monitor and to shield the high voltage. We assume the thickness of the front glass between 3-5 mm.

The tracking system is a touch sensitive film stretched on a 6 mm glass plane, which is mounted 2.6 mm in front of the display plane (Figure 2 (b)). Hence, the overall offset between interaction plane and display plane is 10 mm. The resistive system is capable of tracking a set of passive tangible user interfaces [15] like pens and an eraser, as well as the user's finger. The tracking position is forwarded to the operating system as standard pointing device coordinates. The tracking system was statically calibrated with 6 sampling points. To avoid the influence of a changing viewpoint relative to the target, the calibration is done orthogonally by hand. This implies an inherent static error of the tracking system.

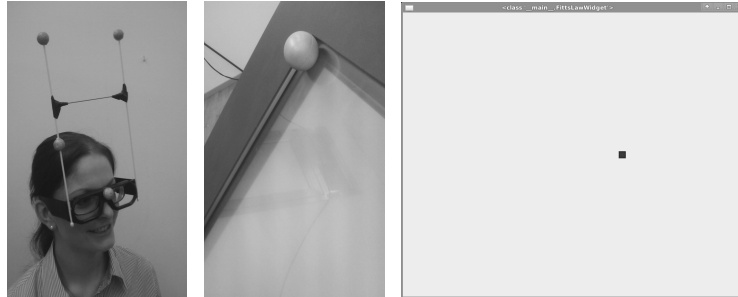


Figure 3. (a) Head tracking module, (b) passive marker on the interactive screen, (c) Interaction application.

### 3.1.4 Tracking System

We used the Qualisys Motion Tracking System with 4 Oqus 300 cameras to continuously track the position of the user's head and of the interactive screen. Each camera emits infrared light, which is reflected by marker balls. After calibrating the tracking system with an error standard deviation of 0.87 mm, the Qualisys tracking software provides 3D positions of the markers' centers. Together with the touch sensitive overlay on the screen, the overall setup provides the 3D coordinates of the viewpoint, the interaction point, and the target point relative to the upper left corner of the display. The viewpoint body consists of 4 markers applied to glasses the user has to wear during the tests. It was also possible to wear additional correcting glasses (Figure 3 (a)). The position of the user's viewpoint is unequivocally defined by the position of the ball, which is mounted next to his nose.

The target position as well as the touch position on the interactive screen is given as display coordinates. As we will discuss in more detail, the hit point and the target position on the screen are provided by the test application as two-dimensional display data. To compare this data with the 3D viewpoint position, the display data is transformed into 3D data. In order to map the display data to 3D spatial data, the display's 3D position also has to be tracked. For doing so, we equipped the display with three markers in the corners (Figure 3 (b)). Since the corresponding vectors are aligned to the display dimensions, the display coordinates can easily be transformed to special coordinates using the display's pixel pitch as dilation factor.

The tracking setup covers the working area in front of the interactive display. Since the system interaction on the interactive screen is limited by the user's arm length (between 662 and 787 (616 and 762) mm of human males (females) [16]), the tracking area is limited to two times the maximum arm length plus the display width times the arm length (distance to the display) times the user's height (between 1629 and 1841 (1510 and 1725) mm of males (females)). The tracking system captures the user's position with an update frequency of 50 Hz.

### 3.1.5 Data Consolidation

As described before, the tracking system provides 3D data from the user's viewpoint and the corners of the interactive screen in mm, while the test software provides the target position and the user's click position in 2D display coordinates. Thus, this different information has to be consolidated to have a single data model.

Based on the pixel pitch of the display and the 3D positions of the display corners, the logging software calculates the 3D position of the 2D display data. The display coordinates are shifted orthogonally to the display plane by the parallax offset of 10 mm with additional respect to the marker's diameter. Given these parameters, the consolidation module queries the 3D tracking software for the markers' 3D positions, triggered by an interaction of the user within 50 mm distance of the target to ignore accidental integrations. Then, the software records the timestamp, the calculated 3D data of viewpoint, as well as target and hit point

relative to the upper left corner of the display. Logging the viewpoint, the interaction point and the target on the screen separately, the setup is robust against movement of the devices.

## 3.2 Measurement Results

Based on the 3D measurement of the touch position on the interaction plane, the target position on the image plane and the user's viewpoint position in front of the screen, we analyze the pointing error of the hardware setup, defined as offset between interaction position and target center for the vertical and horizontal dimension according to the display. Moreover, we deduce the assumed interaction point from the target position and the viewpoint, to show the impact of the parallax error onto the overall interaction error.

In preparation of the observation model, which we shall describe next, we investigate the correlation of the viewpoint and the pointing error.

### 3.2.1 Pointing Error

This section contains results of the pointing error measurements. Due to the small and symmetric target, we assume that the targets are hit in the center. The first aspect is the overall pointing error, which is defined as offset between interaction point and target center. The second aspect shows the influence of the parallax distortion on the overall pointing error, measuring the offset between the actual interaction point and the assumed interaction point. The assumed interaction point is defined on the line between target and viewpoint (Figure 4), which implies not to be affected by the parallax error. Hence, the deviation of the overall pointing error (first aspect) and the offset between assumed and actual hit point (second aspect) indicates the influence of the parallax error on the overall pointing error, with respect to the specific test setup.

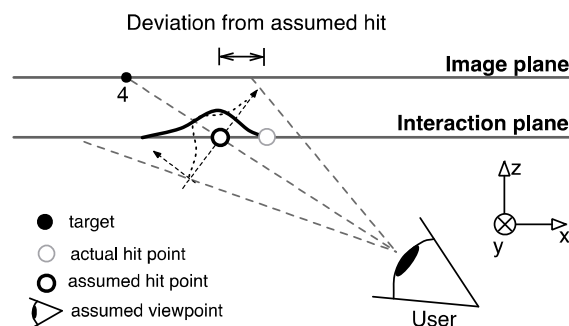


Figure 4. Schematic model of the assumed hit point.

We deduce the overall pointing error directly from the empirical data and analyze two aspects: The pointing accuracy is a measure of how well the user hits the given target [17]. It is indicated by the mean value of the measurements. The pointing precision, on the other hand, refers to the stability (variance) of the data [17], and is a measure of the quality of repeatability.

Figure 5 shows the vertical (a) and horizontal (b) click error with regard to the center of the target. The measurements show that the user did mostly click left of the target's center with a mean distance of -3 mm. 41 % of the interactions lay outside of the target area (33% left and 8% right). The standard deviation of 11 mm indicates a relatively large uncertainty in horizontal precision. The mean of -7 mm in vertical error indicates that the users mostly clicked above the target. With 10 mm, the standard deviation is similar to the horizontal values, although the target was missed in 45% of the interaction (40% above, 5% below). The histogram of the horizontal pointing error shows two clusters around -5 and 0 mm.

The horizontal deviation of the interaction position from the assumed interaction point is illustrated in Figure 5(b). It has a mean value of -1 mm and a standard deviation of 9.4 mm. Although the two domains are not significantly distinguishable (see Figure 5(c)), the deviation of the actual interaction point from the assumed touch point is lower than the overall pointing error. The vertical measurements show a mean value of -6 mm with a standard deviation of 10 mm. Since 41 % (45 %) of the horizontal (vertical) clicks lay outside the target area, 19 % of the interactions with the target area are recognized correctly.



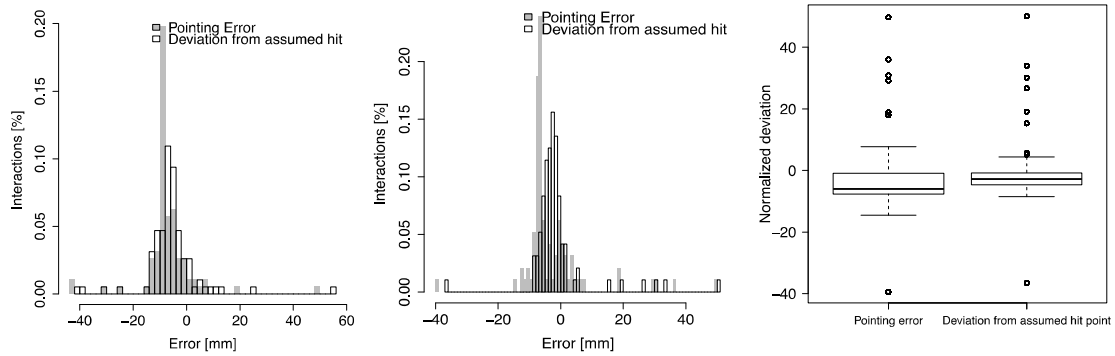


Figure 5. (a) Vertical and (b) horizontal pointing error and deviation from assumed interaction position; (c) boxplot of the horizontal deviation from the target center (pointing error) and the assumed interaction position.

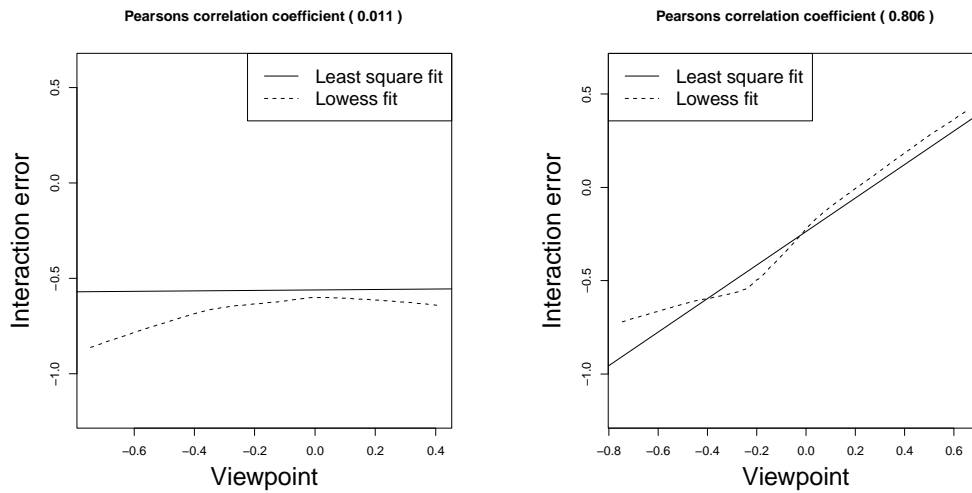


Figure 6. Normalized (a) vertical and (b) horizontal viewpoint and interaction error.

### 3.2.2 Correlation of Pointing Error and Viewpoint

To motivate the dependency of the interaction point and the viewpoint, which is assumed by the parallax correction controller using the pointing error as indication for the viewpoint, we analyzed the correlation between the two domains. As indicated in Figure 2 (c), we investigated the offset of the viewpoint position and the interaction position with respect to the target position. We normalized the offsets with respect to the distance to the display plane and treated the horizontal and vertical dimension separately.

Figure 6 shows scatter plots (with 5505 measurements) of the relation between viewpoint and interaction error in vertical and horizontal dimension. Since Pearson's  $r$  gives a horizontal (vertical) positive correlation coefficient of 0.806 (0.011), the horizontal data is globally positively correlated and the vertical data is not correlated. Hence, we focus on analyzing the horizontal data. The linear model Figure 6 (b) indicates that the correlation is globally biased by a factor of app. 0.3. The LOESS model [24, 34] for local regression shows an almost linear model over the entire space: The correlation is linear: The interaction occurs right (left) to the target if the user is standing right (left) of the target. The area between -0.5 and 0.0 [viewpoint] indicates a correlation of higher order with a medium correlation coefficient of 0.462.

Showing a correlation between the viewpoint position and the interaction error, the measurements motivate to deduce an observation model, which describes the probability of getting an interaction error (observation) given a viewpoint (state). This model will allow the controller to set up the parallax correction based on the viewpoint, which is inferred from the pointing errors.

### 3.3 Discrete Observation Model

Motivated by the measurement results, we now present the discrete observation model for the POMDP parallax correction controller. The model describes the correlation between the interaction error and the user's viewpoint and allows the controller to infer the actual viewpoint from the interaction instead of measuring it directly (with a 3D tracking system). Motivated by the fact, that the correction can only be applied in a discrete way and the interactions are defined with respect to the discrete display coordinates, the final model is formulated as discrete conditional probability distributions over the observation space  $O$  (interaction error) given the state  $s$  (viewpoint) from the state Space  $S$ :  $P(O|S=s)$ .

The correction controller defines the state space as the deviation of the user's actual viewpoint to the left (respectively right) from the assumed viewpoint [9]. The line between the actual assumed viewpoint and the target position defines the expected interaction position (Figure 7). Hence, the offset between expected interaction position and target center defines the actual applied correction of the interaction, so that the interaction on the expected interaction position ( $O_{hit}$ ) would be perfectly corrected to the center of the target. Interacting beside the assumed interaction point indicates a different viewpoint of the user and the controller adapts the assumed viewpoint according to the observation model. If the interaction point around a specific target is within the tolerance interval, the controller assumes that the respective target was aimed for. Therefore, it is mapped to the observation space and emits a discrete observation to the controller. Figure 7 shows three observations:  $O_{hit}$ ,  $O_{left}$  and  $O_{right}$ , which are defined with respect to the inferred viewpoint. Given the actual viewpoint, which is unknown to the controller, the observation model defines the probability of the occurrence of the observation. The controller receives one of the observations and calculates a probability distribution over the viewpoint space applying the Bayes theorem to the conditional probability of the observation model. This new information is used to adapt the assumed viewpoint position and the parallax error correction for the next interaction.

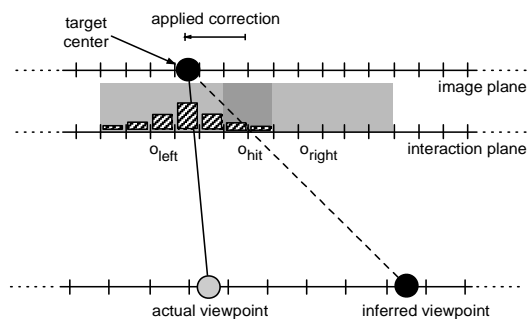


Figure 7. Observation ( $O$ ) model for a given viewpoint and target ( $S$ ) as discrete probability distribution  $P(O|S)$ .

Since the controller uses the interactions (observations) as evidence for the viewpoint (state), we investigate the discriminability of the interaction observations for each viewpoint and show a selection of the resulting conditional probability distributions for the discrete observation space.

First, we discretize the normalized horizontal state space into five intervals with equal length to distinguish between five viewpoints "far left", "slightly left", "above", "slightly right" and "far right" from the target's position ( $[-1.0, -0.55, 0.15, 0.0, 0.15, 0.55, 1.0]$ ). Analyzing the occurring pointing errors for each state separately, we found that the states are distinguishable by the resulting pointing error (see Figure 8(a)). The boxplots show that the areas between the first and the third quartile do not overlap. Although the measurement data does not provide sufficient data for the interval  $[0.55, 1.0]$  (user's viewpoint on the "far left" of the target), due to the few left handed participants of the study, we assume symmetric results. The plots show an increasing variance of the data with increasing offset of the viewpoint.

Due to the significant difference of measurements for the discrete states, the interaction error can be utilized to indicate the viewpoint. As an example for the observation model, Figures 8 (b) and (c) show the discrete probability distribution of two states over five observations.

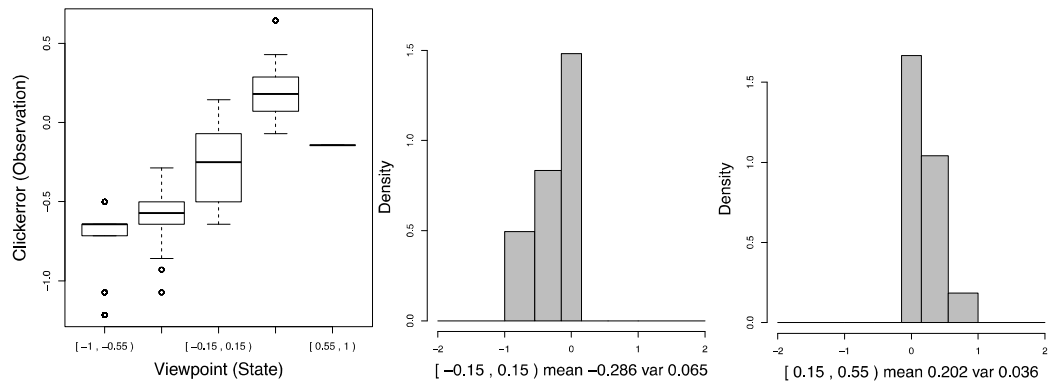


Figure 8. (a) Boxplots of the horizontal click error for 5 discrete viewpoint states. (b),(c) Discrete Observation model for three viewpoints as probability distribution over the observation space:  $P(O|S=s)$

## 4. DISCUSSION

The measurement of the overall pointing accuracy shows a negative mean of the horizontal pointing error. As digital pens were used in the experiments, a likely explanation is the learned behavior of central Europeans to hold pens in a fashion that facilitates writing from left-to-right.

The negative overall vertical pointing error, which means that the user tended to click above the target, could be caused by the fact, that the user's viewpoint was mostly located above the target on the screen, due to the screen setup. The histogram of the horizontal pointing error shows two clusters around -5 and 0 mm. The left cluster indicates the viewpoint to be located left of the target and hit point, and the cluster next to "no error" could be caused by the user trying to correct the parallax error.

The negative mean of the overall pointing error occurs similarly for right- as well as for left-handed people in the vertical and horizontal dimension. This indicates a system inherent error, which could be caused by a dimension-independent error within the static calibration of the tracking system or a wrong offset between interaction plane and display plane. Since the plasma display illuminates gas to plasma, the exact position varies within the cell up to 4 mm. This could also cause the high variation of the pointing error.

The variance of the vertical and horizontal pointing error and dexterity model is similar. The main difference between the vertical and the horizontal data of the pointing error is the mean value, with a distance of 4 mm. Since the dexterity model, which excludes the parallax error, shows a similar distance of 5 mm, without the influence of the parallax error. It could be affected by the different relative position of the viewpoint, under the assumption that the user's motor skills decrease for unfamiliar postures. Analyzing the corresponding data shows that the offset between interaction point and viewpoint is stable for the horizontal direction (low variance), but unstable for the vertical dimension (high variance). Our results confirm the findings in [5]: The user moves mainly in the horizontal direction when working on large interactive screens, while the distance to the screen and the vertical position is stable.

### 4.1 Improvement of the Parallax Error Correction

The quality of an interaction system can be quantified by counting the number of successfully recognized interactions. A correct interaction is defined as hit detection of the target, which occurs if the pointing coordinate is located within the target area. Since the user's goal in the test study was to hit the target area, we interpret every click as an intended click on the target.

In this section, we discuss the interaction improvement using a parallax correction controller without partial observability. This imaginary perfect controller directly measures the viewpoint of the user with a 3D tracking system like we used for the user study. It shifts the interaction position by the offset between assumed interaction point and target center in order to eliminate the parallax error. Although this controller could eliminate the parallax error, it does not reduce the error that stems from the motoric skills and the motivation where to hit the area target. Thus, we can interpret the deviation of the interaction from the

assumed interaction position as click error after viewpoint based parallax error correction. Moreover, the deviation between the resulting target hits indicates the potential improvement of the parallax correction with respect to the specific target (shape and size) and the interaction system. Applying this correction controller, the number of correctly detected pointing interactions would be 43% (82% in the horizontal and 53 % in the vertical dimension). This would increase the pointing quality, since the number of correctly detected interactions from the user study without the correction is only 19%.

## 5. CONCLUSION AND FUTURE WORK

In this paper, we derived a discrete observation model to enable a POMDP parallax correction controller to deduce the user's viewpoint from his interaction. The controller can be implemented into several applications on interactive surfaces like public kiosk applications, cashing and ticketing machines, or large digital whiteboards to eliminate the disturbing parallax error. The increased pointing accuracy allows using smaller pointing targets. This increases the usability of the system, since the number of targets on the screen increases and the number of submenus could be reduced. The main advantage of using a model-based approach is the fact that the controller overcomes the parallax error without any additional and expensive tracking hardware to directly measure the user's viewpoint.

In order to realistically model the user's interaction behavior, we derived the model from empirical data of a user study. We investigated the pointing quality on the test setup and showed the significant impact of the parallax error. Next, we showed the correlation between actual viewpoint and interaction error, which allows the controller to estimate the viewpoint based on the interaction. Since the parallax error depends on the viewpoint, the controller can continuously adapt the correction parameter on the interactive screen to eliminate the parallax error. It does not overcome the error stemming from the motoric skills and the motivation where to hit the area target.

To define an observation model, which is based on the interaction error with respect to a specific target on the screen, the controller has to be able to identify the target, which the user wants to hit. Hence, the targets must be arranged with a certain distance on the screen. We assumed that the target is hit in the center. The model could be extended considering the click positions for different targets. Thus, we will investigate the click position for several targets, which differ in size and shape. This will be deduced from a user study on an interactive screen with minimal offset between the interaction and image plane to minimize the influence of the parallax error. The resulting targeting probability distribution will be used as prior for the estimator to weight the interaction observations.

In a second step, the process model of the user's behavior [5] will be integrated into the estimator to predict the future user position, since the correction has to be done for the upcoming interaction. To set up a non-myopic correction policy with respect to the uncertainty of the estimated viewpoint, the POMDP model has to be planned out. To finally evaluate the estimating controller, we will compare it with a camera based tracking system, which observes the users viewpoint directly.

## REFERENCES

1. Hewett, T.T. et.al., 1992, Curricula for Human-Computer Interaction. *ACM SIGCHI curricula for human-computer interaction*
2. Haller, M. et. al., 2006, The shared design space. *Proceedings of ACM SIGGRAPH 2006 Emerging technologies*, New York, pp. 29
3. Dietz, P., Leigh, D., 2001, DiamondTouch: a multi-user touch technology. *Proceedings of the 14th annual ACM symposium on User interface software and technology*. New York, pp. 226
4. Kunz, A., Fjeld, M., 2010, From Table-Systems to Tabletop: Integrating Technology into Interactive Surfaces. *Müller-Tomfelde, C. (Eds.): Tabletops – Horizontal Interactive Displays*.
5. Migge, B., Kunz, A., 2010, User Model for Predictive Calibration Control on Interactive Screens. *2010 International Conference on Cyberworlds - Cyberworlds 2010*, pp. 32-37
6. Kunz, A., Nescher, T., Küchler, M., 2010, CollaBoard: A Novel Interactive Electronic Whiteboard for Remote Collaboration with People on Content. *2010 International Conference on Cyberworlds - Cyberworlds 2010*, pp. 430-

7. Puterman, M. L., 1994, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Inc., New York
8. Cassandra, A.R., 1998, A survey of POMDP applications. *Working Notes of AAAI 1998 Fall Symposium on Planning with Partially Observable Markov Decision Processes*, pp. 17-24
9. Migge, B., Schmidt, T., Kunz A., 2010, POMDP models for continuous calibration of interactive surfaces. *AAAI Spring Symposium on Embedded Reasoning: Intelligence in Embedded Systems. Technical Report SS-10-04*, pp. 86-92
10. Dao, V.N., Hosoi, K., Sugimoto, M., 2007, A semi-automatic realtime calibration technique for a handheld projector. *Proceedings of the 2007 ACM Symposium on Virtual Reality Software and Technology*, pp. 43-46
11. Lee, J.C., et.al., 2004, Automatic projector calibration with embedded light sensors. *Proceedings of the 17th annual ACM symposium on User Interface Software and Technology*, pp. 123-126
12. Zhang, W., Wu, L., Liu, J., 2009, Research and Application of Matrix Transformation in Position Calibration. *Fourth International Conference on Computer Sciences and Convergence Information Technology*, pp. 933-936
13. Kent, J.C., Warmack, R.J., 2009, Method and apparatus for Laplace constrained touchscreen calibration. *United States Patent Allocation Publication No. 2009/0109195 A1*
14. Hong-Ping, W., Haijiang, G., 2009, Calibration algorithm of touch screen based on BP neural network. *Mechanical & Electrical Engineering Magazine*, vol. 4
15. Ishii, H. et al., 2008, The tangible user interface and its evolution, *Communications of the ACM*, Vol. 51 number 6, pp. 32-36
16. Deutsches Institut für Normung e.V. (German National Standard), 2005, *DIN 33402 Ergonomics - Human body dimensions*
17. eInstruction Global Headquarters Resolution and Accuracy for Interactive Whiteboards, 2007, The full story, [www.einstruction.com](http://www.einstruction.com), Interwrite Learning
18. MacKenzie, I.S., 1992, Fitts' law as a research and design tool in human-computer interaction, *Journal of Human Computer Interaction*, vol. 7, number 1, pp. 91-139
19. ISO, 2000, ISO 9241-9. Ergonomic requirements for office work with visual display terminals (VDTs) – Part 9: Requirements for non-keyboard input devices
20. Tiffin, Joseph; Asher, E. J., 1948, The Purdue Pegboard: norms and studies of reliability and validity. *Journal of Applied Psychology*, Vol 32(3), pp. 234-247
21. Beaton, R.J., Weiman, N., 1984, Effects of touch key size and separation on menu-selection accuracy. *Tektronix Technical Report TR 500-01*, pp. 50-510
22. Wilson, K.S., Inderreiden, M., Liu, S., 1995, A comparison of five user interface devices designed for point-of-sale in the retail industry. *Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting*, pp. 273-277
23. Martin, G.L., 1988, Configuring a numeric keypad for a touch screen. *Ergonomics* 31, pp. 945-953
24. Cleveland, W.S.; Devlin, S.J., 1988, Locally-Weighted Regression: An Approach to Regression Analysis by Local Fitting. *Journal of the American Statistical Association* 83 (403), pp. 596-610
25. Colle, H.A., Hiszem, K.J., 2004, Standing at a kiosk: Effect of key size and spacing on touchscreen numeric keypad performance and user preference. *Ergonomics* 47(13), pp. 1406-1423
26. Beringer, D.B., 1990, Target size, location, sampling point and instructional set: more effects on touch panel operation. *Proceedings of the Human Factors Society 34th Annual Meeting*, pp. 375-379
27. Scott, B., Conzola, M.S., 1997, Designing touch screen numeric keypads: Effects of finger size, key size, and key spacing. *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting*, pp. 360-364
28. Jin, Z. and Plocher, T. and Kiff, L., 2007, Touch screen user interfaces for older adults: button size and spacing. *Universal Access in Human Computer Interaction. Coping with Diversity*, pp. 933-941
29. Kleiss, J. A., 1997, Identifying users' conceptual organization of menu functions in an auto-motive electronic navigation product. *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting*, pp. 944-948
30. Ziefle, M., 2002, The influence of user expertise and phone complexity on performance. *Behavior and Information Technology* 21, pp. 303-311
33. Hofer, R., Kaplan, P., Kunz, A., 2008, MightyTrace: multiuser tracking technology on LC-displays. *Proceedings of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pp. 215-218
34. Cleveland, W.S., 1979, Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association* 74 (368), pp. 829-836.

# EXTENDING A USER OBSERVATION PROTOCOL TO ACCOUNT FOR PSYCHOLOGICAL TRAITS

Yuska P. C. Aguiar<sup>1,4</sup>, Maria de F.Q. Vieira<sup>1,2</sup>, Edith Galy<sup>3</sup>,  
Jean-Marc Mercantini<sup>4</sup> and Charles Santoni<sup>4</sup>

<sup>1</sup>*Universidade Federal da Paraíba, UFCG/LIHM, Depto de Engenharia Elétrica - Campina Grande, Brasil*

<sup>2</sup>*Centre for Excellence in Signal & Image Processing, Dept of Electronic & Electrical Engineering, University of Strathclyde, Scotland, UK*

<sup>3</sup>*Université de Provence (Aix-Marseille I), Département de Psychologie cognitive et expérimentale UFR de Psychologie, Sciences de l'Éducation - Aix-en-Provence, France*

<sup>4</sup>*LSIS, Université Paul Cézanne (Aix-Marseille III), Département Génie Industriel et Informatique, Marseille, France*

## ABSTRACT

Typically the studies of the human error are based on the analysis of incidents and accidents reports. This was also the approach adopted by the authors when analysing human errors in the context of electrical system operation. In a study of a corpus consisting of 31 error reports, extracted over a period of ten years of reports made available by an electricity company in Brazil; it was found that in spite of a detailed account from the system and operational points of view, it was also needed information on the operator's behaviour, in order to understand the error causes. Since this information was only superficially mentioned, in order to complete the study it has been proposed to immerse the user in the situations described in the error reports and to observe and analyze the behaviour when interacting with the system under similar conditions. To achieve this goal, the observation of the user behaviour must be guided by an experimental protocol adequate to behavioural data gathering and analysis. In Psychology, the study of the human behaviour is associated to the study of emotions, through methods and tools based upon the Scherer's Component Model of Emotion. Therefore this work investigates those methods and tools and proposes to extend an existing experimental protocol originally conceived for product usability evaluation, for this purpose. This paper describes the Protocol for Experimental Observation of Interaction (PEOI), and its extended version *E-PEOI* with the addition of the proposed methods and tools. The protocol description encompasses the activities, tools and artefacts necessary to support the observation of the user behaviour during the interaction with electric systems, with the support of Psychology.

## KEYWORDS

Human Behaviour Observation, Registering and measuring Emotions, Human error Studies, Electric systems operation

## 1. INTRODUCTION

Industrial systems are said to be critical when a malfunction can lead to death or serious injury to people, loss or severe damage to equipment, environmental harm or financial loss. Therefore the analysis of accidents and incidents<sup>1</sup> is essential for the human error study, justifying the proposal of preventive strategies such as human interface adaptation, improvement in training programs, and task adaptation. The research on human error focuses on identifying the relation between errors and the human activity that caused it. Thus, contextual factors (task, work environment) and personal factors (operator's profile, emotional state, and behaviour) must be considered when trying to understand the error causes.

Accident report analysis is the path followed by many authors (Rasmussen et al. 1981; van Eekhout & Rouse 1981; Johnson & Rouse 1982). This approach was also adopted when investigating the error causes in accidents reported in the operation of electrical systems at an electricity company in Brazil. The studies were based on a *corpus* of 31 reports of accidents and incidents during a ten year period. The studies resulted in the proposal of: a typology for error scenarios (Mercantini et al. 2004) and, a taxonomy for incident and accident description (Scherer et al. 2010b).

---

<sup>1</sup>Accident: event or situation in which there was some type of loss; Incident: event or situation in which an accident nearly happened.

As already mentioned, the analyzed corpus presented technical information on the error but did not address the operator emotional state and behaviour while performing the task that resulted in the error, needed to understand the human error. To complete the study it has been proposed to observe the operator while performing the same task under similar conditions. Since due to safety regulations it is not possible, to accomplish this observation in a real working environment, this work proposes to observe the operator working activity immersed in scenarios which replicate situations described in the analysed corpus. To support the error studies, the observation must be guided by an experimental protocol (a set of procedures, and tools) that supports the planning, data gathering and analysis, related to the user behaviour.

The Protocol for Experimental Observation of Interaction (PEOI) (Aguar & Vieira 2009) was conceived to support product usability evaluation, and is based on recommendations available in the literature (Preece et al. 2007; Mayhew 1999; Nielsen 1994; Redish 2007). In spite of being functionally acceptable, product users may experience difficulties when using a product or system which user interface does not reflect their characteristics and needs. User satisfaction can be expressed by metrics such as easy to learn and easy to use, which are at the centre of a usability evaluation procedure. PEOI has been employed to support product evaluation in different environments such as laboratory tests, field tests and *in situ*, and with products of different natures. However, PEOI does not clearly address the observation and analysis of the human behaviour.

So this paper presents the extension of this protocol with existing methods and tools found in Psychology that enable the observation of the operator user behaviour though the measuring of emotional states, in accordance with the Scherer's Components Model of Emotion (CME) (Scherer 2001). The CME considers emotion as an episode of interrelated synchronized state changes of subsystems as a function of the evaluation of an external or internal stimulus event. Its components are: cognitive appraisal; physiological reactions (bodily symptoms); behaviour tendencies; motor expression (facial and vocal expression); and subjective feelings (emotional experience). The complimentary data gathered with the support of the extended protocol should allow for the understanding of the human error in the analysed corpus.

This paper is organized in five sections. Section 2 presents the original experimental protocol PEOI. Section 3 introduces the approach employed in Psychology studies to understanding human emotions. Section 4 presents the set of selected tools to be employed in the behavioural data gathering and analysis. Section 5 presents the extended protocol on the basis of Psychology. Finally the conclusion Section presents some considerations and future directions for this research work.

## **2. PROTOCOL FOR EXPERIMENTAL OBSERVATION OF INTERACTION - PEOI**

The original experimental protocol is organized in six steps each of which consists of a process detailed in a set of activities. Each step is associated with one or more objectives to be achieved after execution. The completion of a step is achieved by executing a set of processes and activities. Executing activities is directly related to creating or updating artefacts. Producing artefacts registers the experiment for future reference. PEOI also defines a set of roles to be played by those involved in the experiment. A role defines the responsibilities and expected behaviour during the application of the protocol.

It follows the description of PEOI's steps: Step 1: Planning the experiment consists of characterizing the product/system, its context of use and its users. Step 2: Training when needed, consists on introducing the evaluation team and test participants with the product and its context of use. Step 3: Preparing and Validating the experiment: consists on structuring the experiment and developing the necessary supporting materials. Validating consists in performing a pilot experiment, with a recruited participant. Step 4: Conducting the experiment and Data gathering: implements the experiment's plan, and obtains the data sample. Step 5: Data Tabulation and Analysis: structures and organizes the gathered data for analysis. Produces a diagnostic view for the product based on the user interaction. Step 6 - Presentation of Results: specifies the form, content and media to report the experiment and its results.

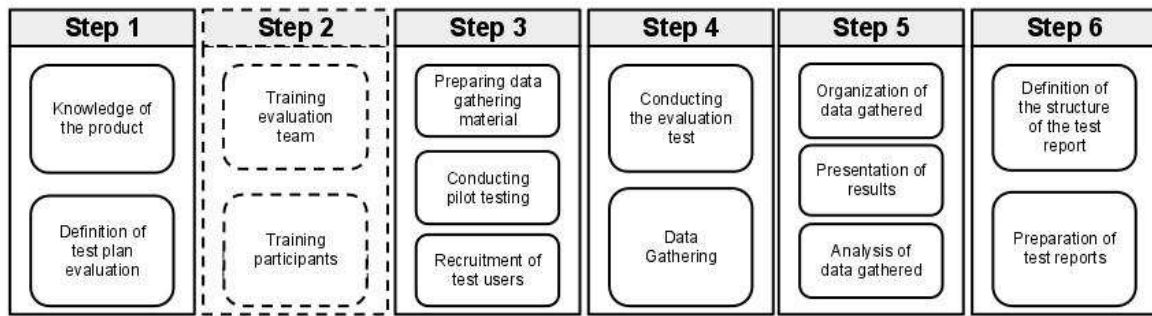


Figure 1. Steps and Processes in PEOI

To understand the behaviour when a user performs a task it is necessary to identify precisely which variables to observe, and determine how and when they should be measured. PEOI employs a set of four methods for data gathering (observation-O, interview-I, questionnaire-Q, document analysis-DA). The data is grouped into the following categories. (I) **General data** is gathered through interview, and aims to clarify the experiment’s objectives. On the other hand, all four methods are employed to gather data on the task, product and context of use. (II) **Pre-interaction**: except for observation, all four methods are employed to gather data on the user profile (personal, professional and contextual). (III) **During-Interaction**: observation is used to gather data on subjective indicators. Here, all four methods, except document analysis, are employed to collect objective indicators on the user activity. In the final step, (IV) **Post-interaction**, all methods (except document analysis) are employed to gather data on the user satisfaction with the product/ system under evaluation. Table 1 presents the relationship between the data categories the data gathering methods.

PEOI was employed in the evaluation of the usability of different products and systems, including the electric system simulator developed to support this study. The objectives and subjective indicators proved to be sufficient from the usability point of view however, regarding the specific interest in understanding the user behaviour it would not be appropriate. Besides the operator profile: (a) physical (manual skills, weight, height), (b) psychological (motivation, attention, temperament, cognitive functions) and (c) clinical (memory, perception, psycho-motor, mental/language functions), it is necessary to know the user emotional state, before, during and after the task performance. For this purpose, besides the methods already employed by PEOI, other methods and tools are required, which are supported in theory and practice by Psychology.

Table 1. Elements of interest to observe along the interaction and methods for data collection – PEOI

	Elements of Interest	Examples	Generic Methods	
I	Objective Assessment	Product release of a new version, Product adaptation,...	Interview	
	Task	Characteristics; work Conditions, time of day, working shift.	All	
	Product and context of use	Auxiliary equipment, environment conditions,		
II	User (profile)	Personal	Age, gender, marital status	All (except O)
		Professional	Job-position, experience in the task, , training level	
		Contextual	Working hours, group work or individual work, task rhythm,	
III	Indicators	Objective	Task duration, help accesses , incorrect actions,	Observation
		Subjective	Ease of understanding and learning, ease of use,	All (except DA)
	User	Satisfaction	During: navigation, help, documentation, product features,	
	Results achieved		Tasks completed, unfinished, abandoned, performed with errors	Observation
IV	User	Satisfaction	After: navigation, help, documentation, product features,	All (except DA)



### 3. PSYCHOLOGY AND THE STUDY OF EMOTIONS

With the advancement and popularization of interactive technologies, the users' emotional state in the interactive systems context has become a valuable source of information to improve the interaction mechanisms offered by the system (Tayary & Le Thanh 2009). Studying human reactions to emotional episodes allows understanding the human behaviour. In this sense, the concept of emotion, the identification of basic emotions (joy, sadness, disgust, fear, anger, and surprise - universal, innate and not reducible) and complex emotions (possible combinations of basic emotions to form complex emotional states), as well as the representation and measurement of emotions, have been areas of interest in Psychology.

The work related to interaction observation in the context of dynamical systems was used as a basis for selecting the variables and related measuring methods adopted in Psychology. In dynamic systems, operator actions are combined with process evolution without operator intervention. Dynamic systems can be classified as: process control, traffic control, disaster management and resources management. Emotions have been studied in dynamic environment context, and in particular in air traffic control, was used as reference for the current study.

The air traffic controllers (ATC) activities focuses on sequencing aircraft approach and maintaining the airspace ordered to avoid collisions (Wickens et al.1997). Into this context, the understanding of controllers' performance and emotional state is a key aspect to promote safety and security in work performance. In general (Cariou et al. 2008; Collet et al. 2009) the emotional variables in the context of ATC are: alertness; tension/stress and fatigue/tiredness. The principal factors that impact on those variables are related to time (time of day, time of service, shift-work, etc.) and workload (intensity of task, number of tasks, complexity of task, concurrency of tasks, etc.), and results on the performance modification.

In the electric system operation context, the analyses of the corpus of reports identified many different factors that may contribute to human errors (Scherer et al. 2010b). The results classified factors in four categories: organization, task, action and operator, with the elements distributed in groups: equipment, material, environment, programming, execution, organizational data, user data, work characteristics (labour), psychological profile, mechanisms of human dysfunction and internal human dysfunction. For the category "operator" and within the group "work characteristics" there is the *state of operator* with the following emotions cited in the reports: anxiety, fatigue/tiredness, confusion, distraction/carefree/indifferent, overexcited/ecstasy, discouraged, tense/stressed, alert (excess), self-confident (excess) and fearful. The emotions mentioned in the ATC's study are also present on this study. Other variable that can be of interest are irritation and satisfaction.

Therefore, the emotions of interest to this study are the basic emotions and those mentioned in the electricity company reports and in the ATC's studies. According to dictionary GALC (Scherer 2005) the terms adopted and their corresponding terms (within brackets) are: joy (overexcited), sadness, disgust, fear, anger, surprise, anxiety, hope (confident), interest/enthusiasm (alertness), relaxation/serenity (distraction/carefree), tension/stress, irritation, boredom (indifferent), contentment (satisfaction). Although there is no mention in GALC to the variables: fatigue/tiredness, confused and discouraged.

### 4. PSYCHOLOGY TOOLS FOR MEASURING EMOTIONS

In the following studies Mahlke (Mahlke & Lindgaard 2007; Mahlke et al. 2006) identified how usability and emotional reactions can determine the user's overall appraisal of the system and thus influence the future decisions and behaviour. In the work it was used CME to structure a range of relevant emotion-measuring methods. Therefore this work is taken as the basis when investigating emotions and for the adaptation of PEOI. It follows a brief description of CME and the list of emotions that can be measured with the respective tools.

Cognitive appraisal is defined as a quick evaluation of a situation that can direct emotional responses (positive or negative). Demir (2009) proposes the following set of appraisal components: consistency of motives, intrinsic pleasure, expectation confirmation, standard conformance, agency, coping potential, and certainty. The tool Geneva Appraisal Questionnaire - **GAQ** assesses the result of an individual's appraisal in the case of a specific emotional episode. GAQ aims to measure: intrinsic pleasantness, novelty, goal/need conduciveness, coping potential and norm/self-compatibility.

**Physiological reactions:** these can be expressed in: cardiovascular, electro-dermal and respiratory measures. Kreibig (2010) presents a review on the investigation of different emotions using a range of emotional induction paradigms. The review argues that the elements most often investigated are distributed in three categories: (i) cardiovascular measurements: heart rate (**HR**), systolic and diastolic blood pressure (**SBP**) and heart rate variability (**HRV**); (ii) respiratory measurements: respiration rate (**RR**); and (iii) electro-dermal measurements: skin conductance level (**SCL**).

**Motor expressions:** these are classed as postural, vocal and facial expressions. Considering the extent and complexity of these data, this work currently focuses only in facial expressions. To measure emotional facial expression, it is necessary to identify the facial expressions of interest and then correlate those with the appropriated emotion (or set of emotions). The most consecrated work in this context was developed by Ekman & Friesen (1978), which resulted in the Facial Action Coding System (**FACS**), but its adoption requires a highly skilled professional. To simplify the process, this work adopts a system for automatic facial expression analysis in real time, the **FaceReader** (Uyl et al. 2005).

**Subjective feelings:** reflects a unique experience consisting of mental and bodily feelings during a particular event. Scherer (2005) claims that no objective method for measuring the subjective experience exists. To access it one must ask the individual to report on his/ her experience ("How do you feel/ felt?"). The Self-Assessment Manikin – **SAM** (Lang et al 1993) is a non-verbal scale, using schematic manikins (icons) to represent different feelings. SAM investigates: valence, arousal, and dominance dimensions. Whereas the Activation-Deactivation Adjective Check List - **AD-ACL** (Thayer 1978) is a multidimensional tool to investigate various transitory arousal states; and considers four sub-scales to measure the relation between energetic and tense arousal. The Geneva Emotion Wheel – **GEW** (Scherer 2005) is a verbal self-report instrument in which the participant is asked to indicate the emotional intensity for a single emotion (or a blend of several emotions) on 20 distinct emotion families (with five degrees of intensity). The **EmotAIX** (Piolat & Bannour 2009) is a dictionary which automatically identifies, categorizes and records the vocabulary of emotion from oral or written language. This vocabulary (literally and figuratively) concerns: emotions, feelings, moods, humour, personality and emotional temperament.

**Behaviour tendencies:** This component is the least explored in the reviewed literature. The studies mention measuring this component through quantitative indicators such as: task completion and time spent on task, which were already contemplated in the original PEOI, therefore, it will not be explored.

Table 2. Relation between measuring tools and observed emotions

CME	Tools	Emotions														
		Found in the reviewed literature										Basic				
		Interest	Anxiety	Hope	Boredom	Relaxation	Fatigue	Irritation	Contentment	Tension	Joy	Sadness	Disgust	Fear	Anger	Surprise
Cognitive appraisal	<i>GAQ</i>		X					X	X		X	X	X	X	X	X
Physiological reactions	<i>Cardiovascular</i>		X						X		X			X		
			X						X		X			X		
		<i>Respiratory</i>		X						X		X			X	
	<i>Electrodermal</i>		X						X		X			X		
Motor expression	<i>FaceReader</i>											X	X	X	X	X
Subjective feelings	<i>SAM</i>		X	X	X	X		X	X							
	<i>AD-ACL</i>	X						X	X		X			X		
	<i>GEW</i>	X						X	X		X	X	X	X	X	X
	<i>EmotAix</i>	X	X			X	X	X	X	X				X		

Besides the listed tools, there are two others which are relevant for this study. The Objective and Cognitive Profile of User - **POCUS** (Scherer & Vieira 2008) structures a system user profile in categories: personal, professional, contextual, physical, psychological and clinical. And **NASA-TLX** (Hart & Staveland 1988) employed to measure the mental workload, employing three dimensions: behaviour (effort and performance), task (physical, mental and temporal demands), and subjective (frustration). Table 2 relates the tools with the emotions of interest for the operator behaviour analysis.

The referred tools are suited to measuring other emotions which were not mentioned because they do not concern this study, such as: pride, feeling love, guilt and jealousy. Nonetheless, other variables can be of future interest, such as: negative and positive appraisal (GAC), and neutral (GEW/FaceReader). As for the variables: confused and discouraged, they were not considered because all the tools already focus on these variables (it is possible to include questions about these emotions in GAQ or additional items in GEW).

## 5. ADAPTATIONS IN PEOI

In spite of extending the protocol's application to support user behavior observation, no changes were required in its general structure, with its steps and processes remaining the same. The changes required were: extending the range of data to be gathered, adding new methods for data collection, and including new activities to be performed by the evaluation team during the experiment, with respective artifacts.

Given the new aspects of interest in the pre-interaction step, POCUS is proposed when gathering data on the user profile. The extended profile is relevant to support identifying personality and temperament traits. Whereas, during the steps: interaction and post-interaction, the focus becomes data gathering on the operator's emotional state. In these two steps the elements of interest are the following emotions: joy, sadness, disgust, fear, anger, surprise, anxiety, hope, interest/enthusiasm, relaxation/serenity, tension/stress, irritation, boredom, contentment, fatigue/tiredness, confused, discouraged, negative, positive and neutral.

Concerning the methods for data gathering, four new groups are being proposed: a) physiological measurements to gather physiological reactions; b) face recognition to gather motor expressions; c) self-report and dictionary to gather subjective feelings. Regarding the team activities during the experiment, the changes concern the measurements to be performed along the process of behaviour observation.

The processes related to: experiment planning activity, executing the plan and data gathering; suffered the highest impact due to the protocol extension. In the process Preparing data gathering material, four activities were modified: (1) defining which data to gather in order to include variables related to the operator's emotional state; (2) including tools for data gathering: cognitive appraisal (GAQ), physiological measures, motor expressions (FaceReader), subjective feelings (SAM/AD-ACL/GEW/EmotAix), user profile (POCUS) and workload (NASA-TLX); (3) specifying the tools and resources required to collect physiological reactions (HR, HRV, SBP, RR, SCL); (4) preparing the artefacts required to perform the experiment: questionnaire, forms/cards, self-report, etc.

In the process Data Gathering, three activities were modified: (5) Pre-test activities: apply POCUS while measuring physiological variables. Those variable values will be used as a reference to be later confronted with the values collected during task activity; (6) Conduct the observation: measure physiological variables (HR/HRV/SBP/RR/SCL), measure motor expressions (FaceReader) and measure subjective feelings (SAM/AD-ACL); (7) Conduct post-test activities: measure cognitive appraisal (GAQ), measure subjective feelings (GEW/EmotAix), measure the workload (NASA-TLX).

In Step 5, the process Analysis of data gathered reflects all the changes introduced in the previous steps. Therefore the data gathered in Step 4 will impact in the analyses process because it will require the correlation analysis between subjective and objective indicators. Not all the tools and data types included in the extended protocol must necessarily be adopted in every experiment. The choice of data depends on the specific aim of the observation, with a specific experiment only encompassing a subset of the human behaviour related variables to be observed (subset of emotions).

The addition of tools to support PEOI-E application implies elaborating new artefacts related to the instantiation of questionnaires and self-reports and to the recording of facial expressions and physiological parameters. However, due to this article's restriction on length, artefact and templates will not be detailed. Due to the specificity of the data to be collected and to analysed, it is recommended to include a psychologist in the team responsible for the protocol application.

## 6. CONCLUSION

This paper argued for the need to understand accidents and incidents in the industrial sector from the user behaviour point of view. The usual approach to human error study, that is report analysis, usually lacks

relevant information on the operator behaviour (emotional state) during risk situations that can lead into error. In the research context in which this work is inserted, that is accidents and incidents in the electricity industry, it is proposed to acquire this information through the observation of operators interacting with a simulated environment capable of reproducing error scenarios described in the reports. In order to collect such specific data, an experimental protocol (PEOI) has been extended (PEOI-*E*) with a set of activities and tools that enable to collect data on the human behaviour through the analysis of emotional state variables.

The set of variables of interest was extracted from studies in two different contexts: electrical systems operation and air-traffic control. Both cases relate to critical systems and dynamic environments. The variables relevant to this study are: joy, sadness, disgust, fear, anger, surprise, anxiety, hope, interest/enthusiasm, relaxation/serenity, tension/stress, irritation, boredom, contentment, fatigue/tiredness, confused, discouraged, negative, positive and neutral. To measure the emotions a set of tools was proposed and grouped according to CME into the categories: physiological reactions (cardiovascular/respiratory/electro-dermal); motor expressions (FaceReader); subjective feelings (EmotAix/SAM/AD-ACL/GEW); cognitive appraisal (GAQ). The changes in the original protocol consisted in including new activities to be performed; corresponding artefacts required; the actors and respective roles.

The extended protocol (PEOI-*E*) is currently being employed in the observation of operators of a system developed to support the elaboration of contingency plans for maritime accidents involving cargo ships containing polluting materials. In the context of electrical systems error studies the research will progress into performing the experiments, replicating error scenarios and gathering the data from the operator behaviour during risk situations; in order to complement the analysis of the corpus of reports. As a consequence of this study it is intended to apply the knowledge on user behaviour in the refinement of a programmed user behaviour model (Scherer et al. 2010a; Ademar Netto et al. 2009) that was conceived to support the development of more ergonomic human interfaces. This model is currently used by designers to simulate user behaviour and analyse the corresponding outcomes to the task being performed by the user. This model is associated to the Method for Conception of Ergonomic Interfaces - MCIE (Turnell 2004).

## REFERENCES

- Ademar V. S. Netto; Aguiar, Y. P. C.; Scherer, D. ; Vieira, M. F. Q. (2009) Context analysis during task execution: an operator model. In: IADIS Multi Conference On Computer Science and Information Systems, Algarve. Proceedings of Interfaces and Human Computer Interaction 2009 and Game and Entertainment Technologies, p. 113-120.
- Aguiar, Y. P. C.; Vieira, M. F. Q. (2009) Proposal of a protocol to support product usability evaluation. In: Fourth IASTED International Conference Human-Computer Interaction, St. Thomas. Proceedings of Fourth IASTED International Conference Human-Computer Interaction, p. 282-289.
- Cariou, M., Galy, E., Mélan, C. (2008). Differential 24-h variations of alertness and subjective tension in process controllers: investigation of a relationship with body temperature and heart rate. *Chronobiology International*, 25:4, p. 597-607.
- Collet, C., Averty, P., Dittmar, A. (2009). Autonomic nervous system and subjective ratings of strain in air-traffic control. *Applied Ergonomics*, 40(1), 23-32.
- Demir, E., Desmet P.M.A., and Hekkert, P. (2009) Appraisal Patterns of Emotions in Human-Product Interaction. *Int. J. of Design*, 3, 2 , p. 41-51.
- Ekman & Friesen (1978) Facial action coding system: A technique for the measurement of facial movement. Palo Alto, Ca.: Consulting Psychologists Press.
- GAQ - Geneva Appraisal Questionnaire available under Research Tools at <http://www.affective-sciences.org/researchmaterial>
- Hart, S. G. & Staveland, L. E. (1988) Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock and N. Meshkati (Eds.) *Human Mental Workload*. Amsterdam: North Holland Press.
- Kreibig, S.D. (2010) Autonomic nervous system activity in emotion: a review, *Biological Psychology* 84, p. 394–421.
- Lang, P. J., Greenwald, M. K., Bradley, M. M., and Hamm (1993) A. O. Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, 30 , p. 261-273.
- Mahlke, S. & Lindgaard, G. (2007). Emotional Experiences and Quality Perceptions of Interactive Products. In J. Jacko (Ed.) *Human-Computer Interaction, Part I, HCII, LNCS 4550* (p.164-173). Berlin: Springer.

- Mahlke, S., Minge, M., Thüring, M. (2006) Measuring multiple components of emotions in interactive contexts, CHI extended abstracts on Human factors in computing systems, Montréal, Québec, Canada
- Mayhew, D. J. (1999) The usability engineering lifecycle: a practitioner's handbook for user interface design, San Francisco, CA: Morgan Kaufmann Publishers Inc..
- Mercantini J.-M., Turnell M.F.Q.V, Guerrero C.V.S, Chouraqui E., Fatima F.A.Q and Pereira M.R.B (2004). Human centred modelling of incident scenarios. IEEE SMC, Proceedings of the International Conference on Systems, Man & Cybernetics, p. 893-898. October 10-13, The Hague, The Netherlands.
- Nielsen, J. (1994) Usability Engineering. Morgan Kaufmann
- Piolat, A., & Bannour, R. (2009). EMOTAIX : Un Scénario de Tropes pour l'identification automatisée du lexique émotionnel et affectif. L'Année Psychologique, 109, p. 657-700
- Preece, J., Rogers, Y. & Sharp, H. (2007) Interaction Design: Beyond Human-Computer Interaction. 2nd Edition.
- Rasmussen, J., O. M. Pedersen, A. Carmino, M. Griffon, and P. Gagnolet. (1981) Classification system for reporting events involving human malfunctions. RISO National Laboratory, Denmark: RISO National Laboratory.
- Redish, G. (2007) Expanding Usability Testing to Evaluate Complex Systems. Journal of Usability Studies, Volume 2, Issue 3, May , p. 102-111
- Scherer K. R. (2001) Appraisal Considered as a Process of Multi-Level Sequential Checking; in K. R. Scherer, A. Schorr and T. Johnstone (eds) Appraisal Processes in Emotion: Theory, Methods, Research, p. 92-120. New York and Oxford: Oxford University Press.
- Scherer, D. ; Ademar V. S. Netto; Vieira, M. F. Q. ; Aguiar, Y. P. C. (2010a) Programming a user model with data gathered from a user profile. In: ADIS Multi Conference on Computer Science and Information Systems, Freiburg. Proceedings of the IADIS International Conferences Interfaces and Human Computer Interaction and Game and Entertainment Technologies. p. 139-146.
- Scherer, D., da Costa, R.C., Barbosa, J.G., Vieira, M. de F.Q.. (2010b) Taxonomy Proposal for the Description of Accidents and Incidents in Electrical Systems Operation. In: ACM SIGCHI Symposium on Engineering Interactive Computing Systems. Berlin, Germany.
- Scherer, D.; Vieira, M. F. Q. (2008) Accounting for the Human Error when Building the User Profile. In: Third IASTED International Conference Human-Computer Interaction, Innsbruck. Proceedings of Third IASTED International Conference Human-Computer Interaction. Zurich : ACTA Press, v. 1. p. 132-137.
- Scherer, K. R. (2005) What are emotions? And how can they be measured?, Social Science Information 44 (4), p. 695–729.
- Tayari, I., Le Thanh, N (2009) "Modélisation des états émotionnels par un espace vectoriel multidimensionnel" Research Report
- Thayer, R. E. (1978). Toward a psychological theory of multidimensional activation (arousal). Motivation and Emotion, 2, p. 1-34.
- Turnell, M. F. Q. V (2004). Accounting for Human Errors in a Method for the Conception of User Interfaces In: International Mediterranean Modeling MultiConference - I3M'04. Genova, Italy.
- Uyl, M. den, Kuilenburg, H. van & Lebert, E. (2005). FaceReader: an online facial expression recognition system. Measuring Behavior, 5th International Conference on Methods and Techniques in Behavioral Research, Wageningen, The Netherlands.
- van Eekhout, J. M., and W. B. Rouse (1981). Human Errors in Detection, Diagnosis, and Compensations for Failures in the Engine Control Room of a Supertanker. IEEE Transactions on System, Man, and Cybernetics, 12 ed.
- Wickens, C. D., Mavor, A. S., McGee, J. P. (1997). Flight to the future: Human factors in air traffic control. Washington, DC: National Academy Press.

# LAYOUT TEMPLATES — LET USERS RULE USER INTERFACES

Florian Haag, Christiane Taras and Thomas Ertl

*Visualization and Interactive Systems Institute (VIS), University of Stuttgart  
Universitätsstraße 38, 70569 Stuttgart, Germany*

## ABSTRACT

Graphical user interfaces are abundant in end user software. User interfaces with the same purpose but slightly or completely different layouts are used across many different applications. So far, some approaches to let users rearrange these user interfaces have been presented, but they are often very complex and require manual configuration for every single application. In this work, we suggest Layout Templates as a concept that can help rearrange user interfaces across various applications while maintaining a manageable amount of options.

## KEYWORDS

Graphical User Interfaces, Customization, User Interface Layout, End-User Development

## 1. INTRODUCTION

Many software applications these days offer graphical user interfaces. Settings, commands and input facilities are distributed throughout different parts of user interfaces. Each of these sections of a user interface is dedicated to a specific task, with many tasks being alike or overlapping with others. Also, while there is a great variety of computer programs for different purposes, many of them feature similar sub-tasks, such as the process of opening or saving files or setting up a printer.

Just like the rest of the application, user interfaces are designed and developed by software engineers rather than by end users. Only occasionally are means provided to the user to adapt the pre-defined user interface to his or her special requirements and preferences. Where present, though, they are used by a considerable amount of users (more than 75%), as shown in an early study (Page et al., 1996). More than 50% of the users attempt to restructure the user interface in some way. The lack of such features, on the other hand, prevents users from configuring their settings based on their preferences.

In these days, the advent of both big screens and small mobile devices requires applications to run on a multitude of display resolutions and aspect-ratios. This adds to the importance of configurability, especially concerning the user interface layout.

## 2. RELATED WORK

Various attempts have been undertaken to provide the user with the ability to customize the user interfaces of applications that do not provide sufficient customization features on their own. Note that the focus is on modifying existing user interfaces, while means to create user interfaces from scratch, such as with user interface description languages (Nichols & Myers, 2009), are not considered here.

Edwards et al. proposed a system that can apply visual and geometrical transformations to user interfaces or parts thereof (Edwards et al., 1997). Similarly, Stuerzlinger et al. presented the User Interface Façades (Stuerzlinger et al., 2006) that let users rearrange and hide arbitrary regions of a given graphical user interface being displayed. Lutteroth and Weber suggested a system that lets users rearrange user interface elements on a grid layout (Lutteroth & Weber, 2008: 1). Maulsby recommended reordering elements based on the order of actions performed by the user (Maulsby, 1997).

All of these systems focus on adapting single applications, but users may also want to customize certain user interfaces to make them appear more like other applications. Such consistency across a whole working environment is often regarded as a basic requirement to user interfaces (Nielsen, 1989). Even when a general consistency across all applications is not desired, a topic-related consistency that simplifies switching between several tools and tasks in the same topic area is still considered a good idea (Grudin, 1989). To achieve this consistency, Lee et al. suggest in their 2010 work to have user interface elements automatically copy their settings from other elements. This creates a consistency with those other elements. Alternatively, if the user interface is based on an XML description, transformations can be specified with the XSLT language before the user interface is actually rendered (W3C, 2007).

### 3. MOTIVATION

While examining existing solutions, we concluded that the desired degree of customization can be reached in many cases. However, we were dissatisfied with two major aspects: the trade-off between complexity and flexibility, as well as the applicability of a configuration across various applications.

#### 3.1 Complexity vs. Flexibility

The first aspect can be divided into two facets. Existing solutions either provide an unnecessarily restrictive set of options or they are very flexible and hence too complicated to handle.

In the former case, only single attributes of the user interface can be modified. Whether the customization is useful for a particular user depends on whether the developer of the user interface foresaw the most important needs of this user. For example, the development environment SharpDevelop (ICSharpCode Team, 2010) offers a dock widget with various categories of text templates, but only the whole widget can be docked at a certain location; a single category such as the ASCII character table cannot be docked separately. Similarly, reordering interface elements as suggested by Maulsby in 1997 based on the frequency of use allows for flexibility, but does not allow any direct control by the user.

On the other hand, solutions that provide actual flexibility tend to take this as far as possible, providing means to completely restructure the whole user interface. Naturally, this requires an extensive set of instructions to describe all possible transformations. This is particularly evident in the case of XSLT, where users have to familiarize themselves with the XPath language to select nodes to transform. Hence, harnessing the full range of possibilities of such a customization system can be difficult for users. Mackay had earlier suggested in her 1991 study that complicated customization takes up resources that should be spent on work rather than on learning how to adapt tools, thus discouraging users from customizing the software. Another possible drawback of great flexibility is that many actions can lead to undesired results. For example, the User Interface Façades (Stuerzlinger et al., 2006) allow a virtually unlimited reordering of user interface elements, without any guidance based on the internal hierarchy of the original user interface.

#### 3.2 Applicability across Applications

Second, many solutions do not take into consideration that user preferences apply to many similar user interfaces across numerous applications. Different programs offer dialogs with similar or even identical purposes, yet they can rarely be configured system-wide. Instead, their appearance depends on the runtime libraries referenced by the developer or on the (version of the) operating system that is currently being used. An example is shown in figure 1, where a simple user interface element such as a Find dialog box is shown in two quite different variations.

As stated above, a consistency across all applications employed by a user is generally desirable. Differently-styled applications based on different user interface libraries undermine this consistency, as do disconnected configuration facilities in various programs. Instead, if a single configuration utility was provided that allows configuring many applications at the same time, this could improve consistency across programs, as was noted by Nielsen in 1989.



Figure 1. Two different Find dialog layouts (Windows default and Microsoft Word 2007)

### 3.3 Goals

Based on these findings, the goal was to find a way to restructure user interfaces of several applications at a time without requiring any overly complicated procedures. To achieve a desirable degree of flexibility and complexity, a basic scheme for that configuration was required. It would have to work the same for all affected applications. This would ensure a certain degree of simplicity. At the same time, it could change the structure of user interfaces, thereby providing more flexibility than a few single pre-defined options. The fact that the configuration would have to be applied to more than one application, particularly to applications developed with a variety of programming libraries, suggested that an external configuration tool would be convenient. Such a tool could manage the configuration settings and apply them to any running application from the outside.

## 4. LAYOUT TEMPLATES

As a solution to the aforementioned problems, and to reach the defined goals, we have developed a configuration system that allows for a flexible restructuring of graphical user interfaces.

The configuration options are applicable across several applications at a time without having to adjust the settings for every single application. On the technical side, this can be achieved by examining the structure of displayed graphical user interfaces using system APIs, adapting the structure based on the settings and then displaying the restructured user interfaces. From a conceptual point of view, this is because the settings do not contain any references to any specific widgets in a particular user interface. Instead, they try to recognize certain user interface elements based on a set of rules before applying the configuration. These rules are essentially Boolean filter expressions that can compare the type, the caption or any other attribute of user interface elements to expected values. If the comparison returns true, the interface element was recognized and will be manipulated by the configuration system. In many of the test runs conducted during the development, comparisons to check for a given type (“button”, “checkbox”) and display text were sufficient, although more elaborate filters can be applied when needed.

For the configuration process, a three-step approach was chosen. For a graphical depiction of the steps as well as the resulting layout, see figure 2.

- Firstly, a decision regarding which parts of a user interface layout are going to be affected by the configuration system must be made.
- Secondly, the desired user interface layout must be specified (a Layout Template).
- Thirdly, elements found in the original user interface must be linked to the desired layout.

The process is described in a more in-depth way as follows:

For the first step, the user specifies filter rules to select a user interface, or a portion of it. This part of the user interface will be restructured. The configuration system will replace this selected part with a new empty parent element.

In the second step, the user defines a Layout Template. This Layout Template definition specifies how a rectangular area is divided into regions. Each of these regions has a list of category names, which are arbitrary character strings. The user may specify any number of category names for each region. These names define the categories of user interface elements that will be placed in the respective region. Also, a “catch-



all” region that will receive uncategorized elements may be specified. This is useful because not all elements of a user interface are known beforehand and some should simply be included in the new layout as they are.

Finally, in the third step some or all widgets in the considered user interface have to be categorized, so they can be assigned to one or more regions in the Layout Template. To achieve this, the user once more defines filter rules along with category names. Optionally, a priority may be specified with each of these category names as an integer number. Any widget that matches such a defined rule will be added to the categories specified for that rule. Within that category, widgets are sorted by the priority, where different priorities were indicated.

Once the user has executed these three steps of configuration settings, the configuration system can use the settings to restructure the user interface elements. They are inserted into the new parent element (from step one), based on the positions indicated by the regions of the selected Layout Template.

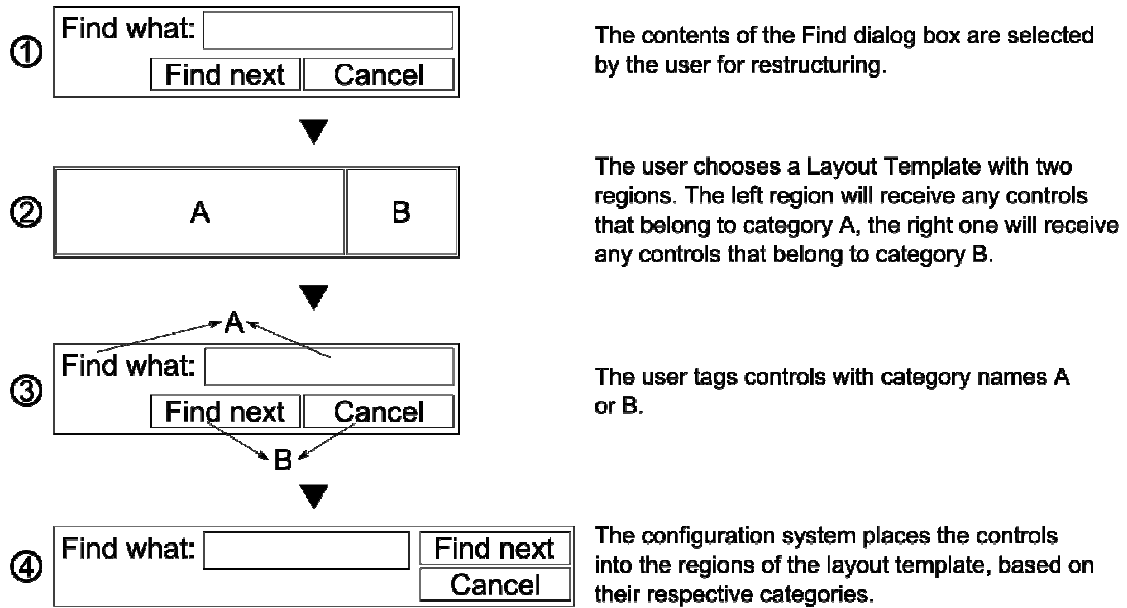


Figure 2. Transforming the layout of (a simplified version of) Microsoft Word's Find dialog to match that of the system's default Find dialog: Three configuration steps and the result

As described above, Layout Templates divide a rectangular area into regions. These regions are aligned based on a basic layout. Upon defining a Layout Template, the user must choose such a basic layout from the set of implemented layouts. Generally, wide-spread layout managers such as a grid layout or a flow layout are offered here, but as an extension, other basic layouts such as the Auckland Layout Model (Lutteroth & Weber, 2008: 2) could also be implemented. To reduce the difficulty of reproducing the Layout Templates concept, the simplified UML class diagram in figure 3 shows how the definition of a Layout Template is structured in memory.

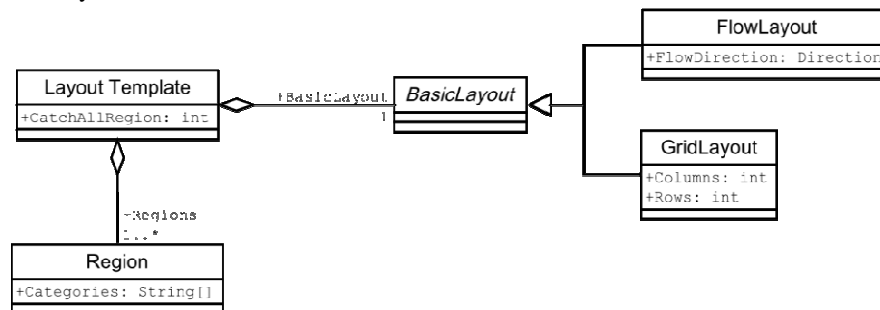


Figure 3. The object structure of a Layout Template definition in a simplified UML class diagram: Each Layout Template refers to a basic layout and defines a set of regions. Basic layouts can have layout-specific attributes.

Depending on the basic layout chosen, some further settings will have to be provided by the user. For example, a grid layout will require some additional information about the number of columns and rows, and a flow layout will need to know whether to flow horizontally or vertically. Other basic layouts, once implemented, may offer other choices. That is why basic layouts are supposed to be derived from a common base class, as shown in figure 3.

Lastly, each region has the aforementioned list of category names. As these can be arbitrary strings and as there may be any number of category names in the list, figure 3 indicates an array of strings. The actual implementation may choose a more suitable dynamic-length list type, depending on the available library.

For the sake of reusability, Layout Templates are usually defined once and can later on be referred to in step two of the configuration process. This results in the settings object structure depicted in the simplified class diagram in figure 4, where Layout Templates are defined on a top level and can be referred to repeatedly.

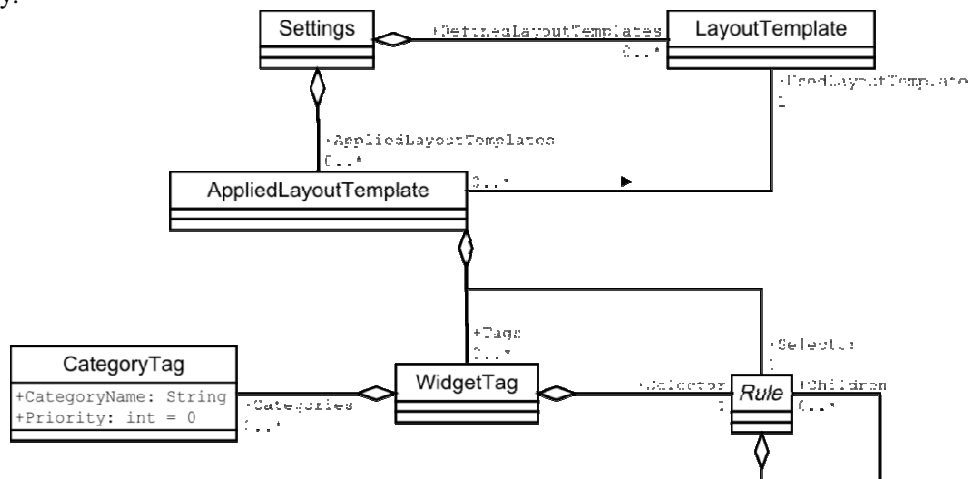


Figure 4. The object structure of the configuration data: When using a Layout Template, its definition is taken from a global list (step 2 in figure 2). The same rule-based filtering mechanism is employed for selecting the user interface part to apply the Layout Template to (step 1) as well as for tagging widgets (step 3).

As can also be seen in figure 4, the same rule-based filtering mechanism is used for the selection of a user interface region to restructure, as well as for the selection of widgets that are to receive category tags. To allow for hierarchically nested expressions as rules, an abstract *Rule* base class that represents a single expression node, which may contain sub-nodes, is depicted here. The root node of such a rule tree must yield a Boolean result value. Details on how to ensure this result value are not provided here as this is an implementation detail that depends on the chosen technology. A *WidgetTag* instance can store any number of category tags that are assigned to the user interface elements for which the selector rule returned true. As pointed out above, specification of a priority along with a category name is not required; in such a case, a default priority will be used.

## 5. IMPLEMENTATION

The current implementation of Layout Templates was created in the context of the HyperBraille project (METEC AG, 2010). That project develops a prototype for a tactile display with a resolution of  $120 \times 60$  dots that can display Braille text, in addition to simple graphical elements. The project aims at transferring existing graphical user interfaces to the tactile display, where they are made configurable with Layout Templates. Originally intended to optimize the graphical user interface for the limited space on the tactile device, it soon became evident that the Layout Templates technique would be useful beyond the HyperBraille project. Since the geometrical structure of the user interfaces is maintained during the conversion to tactile output, the different representation does not have any conceptual impact on the Layout Templates system. Therefore, we chose to present Layout Templates as a concept that is suitable for any kind of geometry-based user interface, to allow more users and more systems to benefit from the concept.

To analyze the graphical user interfaces, the HyperBraille project uses *UI Automation* (UIA) (Kraus et al., 2008). UIA is a programming interface provided by Microsoft Windows to help connect assistive technology with standard graphical user interfaces. Using UIA, state changes are captured to update the tactile output in HyperBraille, and can generally be used as a trigger to apply configuration settings. Likewise, property values can be retrieved using UIA, which can then be evaluated with the filter rules described above. As the Layout Templates concept can be applied not only to the special tactile output of the HyperBraille project, but also to graphical user interfaces, UIA would be used beyond accessibility enhancement to allow a general configuration of user interfaces.

The configuration settings that define the Layout Templates and set the rules to apply the templates to user interfaces are stored in an XML format when serialized to disk. XML was chosen due to the widespread availability of ready-made readers and writers, thus no new parsing engine was required to read the settings. The structure of the XML is closely related to the structure of the settings data, as illustrated in figure 5.

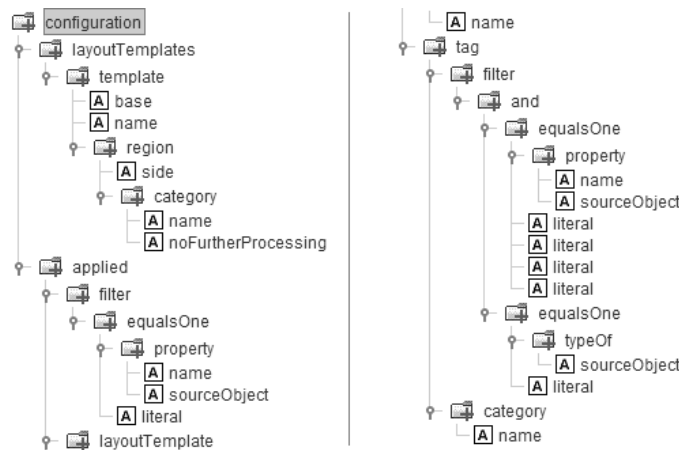


Figure 5. The structure of an XML file that holds configuration data for the Layout Templates system (wrapped in two columns), as displayed in XML visualizer (Chmelar, 2009). In a comparison with figures 3 and 4, the similarity to the memory data structures becomes apparent. For example, a Layout Template (*template* XML node, left column) contains one or more regions (*region* XML node); a *WidgetTag* seen in figure 4 (here: *tag* XML node, right column) has a filter (*filter* XML node) that consists of nested rules (for example *and* and *equalsOne* XML nodes).

While we acknowledge that users without any programming experience may have certain difficulties directly editing XML files, the availability of readers and writers simplifies the future creation of graphical configuration front-ends. Structurally, those front-ends can use the same approach of keeping a list of Layout Template definitions, as well as a list of applied Layout Templates and widget tags.

## 6. HYPERBRAILLE EXAMPLE

As an example, the common elements of the Find dialog layouts from figure 1 should be transformed into the same layout in HyperBraille: A dialog box with buttons on the right-hand side of the window, with *Find Next* above *Cancel* and the dialog contents on the left-hand side. Therefore, the simple Layout Template shown in figure 6 is defined, based on a grid layout with two columns and one row.

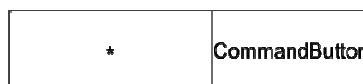


Figure 6. The Layout Template used to restructure the Find dialog

It will be applied to the contents of any Find dialog, hence the title bar is not subject to the Layout Template. The region in the right column will receive any controls categorized as *CommandButton*. Any other controls will be added to the region in the left column.

Now that the Layout Template is defined, the controls have to be tagged. For that purpose, the following filters are defined:

- Any button control with the caption *Find Next* is categorized as a *CommandButton* with priority 2.
- Any button control with the caption *Cancel* is categorized as a *CommandButton* with a priority of 1.
- All other button controls are categorized as a *CommandButton* with a priority of 0.

These rules ensure that all affected Find dialog boxes, such as the one on the left-hand side of figure 7, are displayed like that on the right-hand side of the same figure. Any additional buttons would be below the *Cancel* button. As the text box and its label are not categorized, they end up in area A, as would any other input controls that might be present in some applications. By extending the first filter to also encompass *Ok* buttons, command buttons of a variety of different dialog boxes could be repositioned.

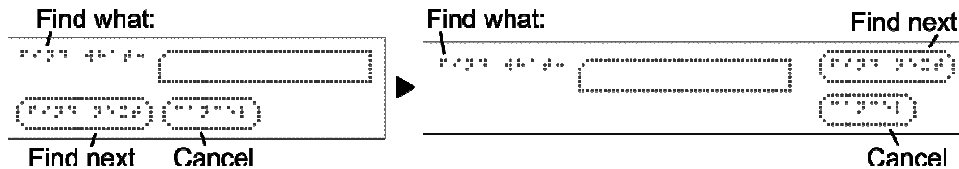


Figure 7. Transforming a Find dialog box using a Layout Template in HyperBraille

## 7. XAML EXAMPLE

As we have pointed out above, the Layout Templates concept can also be applied to ordinary user interfaces. To corroborate this claim, we present a Windows Presentation Formation dialog box specified in the XAML language (figure 8, left). To present a realistic dialog box rather than a contrived example, the layout of the printer setup dialog box from Microsoft Visual Studio 2010 was replicated. By defining some Layout Templates and applying the configuration steps described in section 4 of this paper manually to the XAML source code, the dialog could be transformed into what is shown on the right side of figure 8.

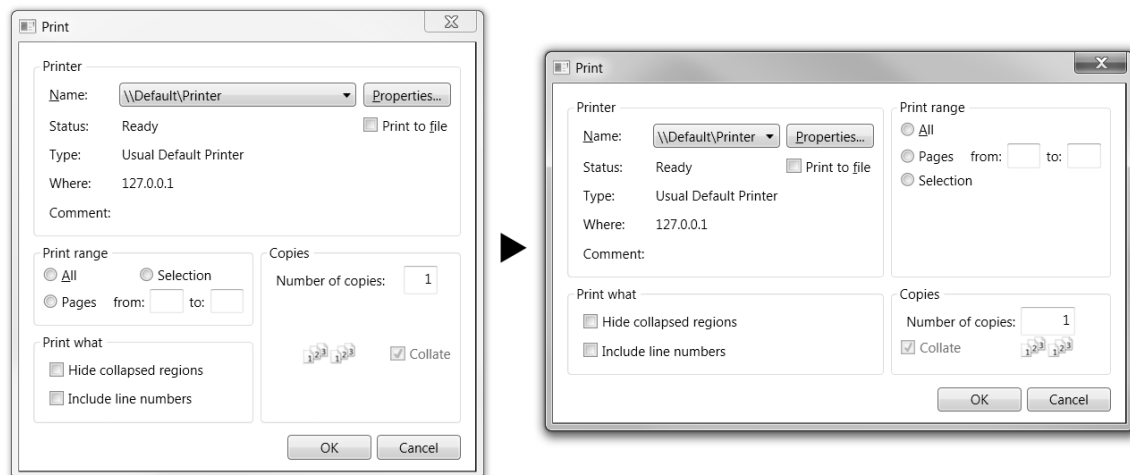


Figure 8. Transforming a Printer Setup dialog box written in XAML

Basically, three grid-based Layout Templates were defined and used:

- One Layout Template replaces the contents of the *Print range* group box. It places the *All* and *Pages* options (the latter along with the page range input elements) present in most applications at the top and adds new lines for any additional application-specific option, such as *Selection*. This change was inspired by figure 2.5 in Wessel's 2002 book about GUI design.

- A second template replaces the contents of the *Copies* group box to bring the check box to the left.

- Finally, a third template replaces the whole dialog box contents to reposition all group boxes.

This shows that the Layout Templates concept can be applied to conventional user interfaces, as well. The actual implementation of a module that would modify XAML GUIs can be achieved in a variety of ways, for example using some dedicated code to modify XML or by generating XSLT code that can be run using an XSLT processor. This is an implementation detail and is not discussed here.

## 8. CONCLUSION AND FUTURE WORK

As we have presented above, Layout Templates are a general concept that can be applied to graphical and other forms of geometry-based user interfaces. This concept can be used to modify and rearrange graphical user interfaces across several applications and provides an extensible set of options. At the same time, it remains simple enough to be described within few lines of text. Therefore, we are convinced that we provide an easy-to-use concept to let users adapt their working environments as they prefer.

As with every concept, the presented system has certain limitations. As mentioned above, its use is currently restricted to restructuring geometry-based user interfaces. However, the Layout Templates concept was designed with extensibility in mind. Even a plugin system that could let users download and add additional base layouts or possibly more transformations or filter rules could be implemented. Future research will show new ways of using Layout Templates to transform user interfaces, in order to adapt the interfaces to a variety of output devices or possibly even to support different interaction paradigms. Moreover, to further simplify editing the configuration, the addition of a graphical configuration editor application is a goal in the ongoing development to further enhance the Layout Templates system.

## ACKNOWLEDGEMENT

This research was partly funded through the HyperBraille project (German Ministry of Economy and Technology (BMWi) grant number 01MT07004).

## REFERENCES

- Chmelar, P., 2009. *XML visualizer*. <http://sourceforge.net/projects/xmlvisualizer>
- Edwards, W.K. et al., 1997. Systematic Output Modification in a 2D User Interface Toolkit. *Proceedings of the 10th annual ACM symposium on User interface software and technology*. New York, USA, pp. 151-158.
- Grudin, J., 1989. The Case Against User Interface Consistency. *Commun. ACM*, Vol. 32, No. 10, pp. 1164-1173.
- ICSharpCode Team, 2010. *Sharpdevelop*. <http://sharpdevelop.com/>
- Kraus, M. et al., 2008: An Off-Screen Model for Tactile Graphical User Interfaces. *Computers Helping People with Special Needs, LNCS*, Vol. 5105, pp. 865-872.
- Lee, L.C. et al., 2010: Improving End-User GUI Customization with Transclusion. *Proceedings of the Thirty-Third Australasian Conference on Computer Science*. Darlinghurst, Australia, pp. 163-172.
- Lutteroth, C. and Weber, G., 2008: End-User GUI Customization. *Proceedings of the 9th ACM SIGCHI New Zealand Chapter's International Conference on Human-Computer Interaction*. New York, USA, pp. 1-8.
- Lutteroth, C. and Weber, G., 2008: Modular Specification of GUI Layout Using Constraints. *Proceedings of the 19th Australian Conference on Software Engineering*. Washington, DC, USA, pp. 300-309.
- Mackay, W.E., 1991: Triggers and Barriers to Customizing Software. *Proceedings of the SIGCHI conference on Human factors in computing systems*. New York, USA, pp. 153-160.
- Maulsby, D., 1997: Inductive Task Modeling for User Interface Customization. *Proceedings of the 2nd international conference on Intelligent user interfaces*. New York, USA, pp. 233-236.
- METEC AG, 2010: *HyperBraille*. <http://www.hyperbraille.de/>
- Nichols, J. and Myers, B.A., 2009: Creating a Lightweight User Interface Description Language: An Overview and Analysis of the Personal Universal Controller Project. *ACM Transactions on Computer-Human Interaction*, Vol. 16, No. 4, article 17.
- Nielsen, J., 1989: Coordinating User Interfaces for Consistency. *SIGCHI Bulletin*, Vol. 20, No. 3, pp. 63-65.
- Page, S.R. et al., 1996: User Customization of aWord Processor. *Proceedings of the SIGCHI conference on Human factors in computing systems*. New York, USA, pp. 340-346.
- Stuerzlinger, W. et al., 2006: User Interface Façades: Towards Fully Adaptable User Interfaces. *Proceedings of the 19th annual ACM symposium on User interface software and technology*. New York, USA, pp. 309-318.
- W3C, 2007: *XSL Transformations (XSLT) Version 2.0*. <http://www.w3.org/TR/xslt20/>
- Wessel, I., 2002: *GUI-Design: Richtlinien zur Gestaltung ergonomischer Windows-Applikationen*. Carl Hanser Verlag, Munich, Germany.

# USABILITY EVALUATION METHOD BY INTEGRATING VARIOUS TYPES OF EVALUATIONS

Junko Shirogane<sup>1</sup>, Yuichiro Yashita<sup>1</sup>, Hajime Iwata<sup>2</sup> and Yoshiaki Fukazawa<sup>2</sup>  
<sup>1</sup>Tokyo Woman's Christian University / <sup>2</sup>Kanagawa Institute & Technology, Waseda University

## ABSTRACT

To ensure software is usable, usability evaluations are commonly performed. However, most evaluations are individually performed, and few methods integrate individual usability evaluations. In our research, we propose a method to automatically integrate various types of usability evaluations. Our method adds functions to record the operation histories of the target software. This information is then used to perform individual usability evaluations focusing on efficiency, errors, and learnability. Individual evaluations are then integrated to identify usability problems.

## KEYWORDS

GUI, usability, usability evaluation, operation history

## 1. INTRODUCTION

To develop usable software, usability of software GUIs are very important, because end-users directly interact with GUIs. For this purpose, there are various methods. For example, GUIs are developed along with usability guidelines and patterns, and developed GUIs are evaluated and improved. Especially, whether end-users feel GUIs usable differs among end-users. Thus, developed GUIs are often evaluated and improved iteratively in terms of usability.

Various usability evaluation methods have proposed, including experimental methods in usability testing [Barnum-2001] and analytic methods in usability inspection [Nielsen-1994 (1)] [Nielsen-1994 (2)], and each method has a different perspective. That is, when several types of usability evaluations are performed for a specific software package, numerous problems are found according to each usability evaluation. Consequently, it is difficult to resolve all problems due to software development costs and schedule. Thus, the identified issues should be prioritized or grouped together to elucidate the more serious problems.

Herein we propose a method to integrate various types of usability evaluations. Our method initially generates functions to record operation histories of end-users in software. This information is then used for usability evaluations. Finally, the results of each usability evaluation are integrated based on the priorities of usability categories [Nielsen, 1994 (1)], which are determined by software developers and end-users.

## 2. FEATURES OF OUR METHOD

Various types of usability evaluations are performed using operation histories by end-users, and the evaluation results are integrated. The features of our method are described below.

### 2.1 Automatic Evaluation of Usability

In our method, our system automatically adds functions to record operation histories of target software using its source programs. In this way, only by operating our system, functions to obtain operation histories can be added to the software. Then our system performs usability evaluations based on these histories. Because software developers and end-users only operate the target software and determine the priorities of the usability categories, the usability evaluation process is mostly automatic.

## 2.2 Prioritization of Usability Problems

Additionally, our method can integrate various types of usability evaluations, allowing usability problems to be prioritized based on the categories determined by software designs and end-users. This will enable software developers to determine which issues to address first. Consequently, the costs and schedule of software development are easy to estimate. Moreover, the usability along with the opinions of software developers and end-users can be easily realized.

## 3. INTEGRATION OF USABILITY EVALUATIONS

Figure 1 depicts a process of our method.

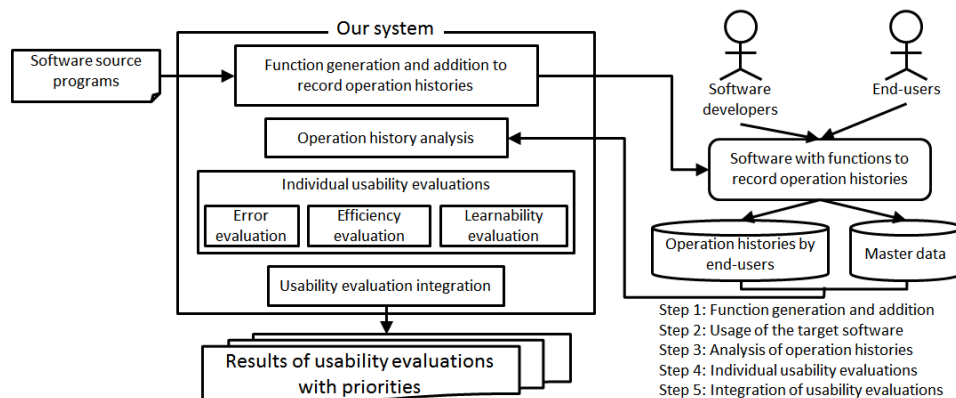


Figure 1. Process of our method

### 3.1 Function Generation and Addition

Functions to record operation histories are generated by analyzing source programs of the target software as programs of AspectJ [Miles-2005] because AspectJ can add additional functions without modifying the original source programs of the software written by the Java programming language. To generate the functions, the following is extracted from the source programs of the target software.

- Variable names of the widgets in a window
- Class names of the widgets in a window
- Superclass name of each class

The generated functions are then added to the target software. Consequently, operating the software, the operation histories are recorded.

### 3.2 Usage of the Target Software

The operation histories are recorded for usability evaluations as the target software with functions to record histories is used. When operating software for usability evaluations, software developers order a certain task, and end-users operate the target software to achieve the task. In this case, two types of operation histories must be recorded: end-user data and master data.

In our method, the usability evaluations are performed by comparing the operation histories of end-users to those of expert users. Operation histories by end-users are called "end-user data", while those by expert users are called "master data". Expert users may include software developers. Ideally end-user data for a task will be comprised of data from many end-users. In contrast, master data requires input from an expert.

### 3.3 Analysis of Operation Histories

Our system analyzes the operation histories, and extracts the following items for a widget or a window:

- Widget coordinates in a window
- Window in which widgets are located
- Window activation
- User event associated with the widget, such as a mouse click and key stroke
- Time that the user events occurs in the widget, and windows are activated

Also, usability evaluations are performed based on windows, as described below. Therefore, operation histories are divided into windows using window activation information. By this division, sequences of window switching are extracted.

## 4. USABILITY EVALUATION

Evaluations are performed using the obtained operation histories. First, individual evaluations are performed, and then the results are integrated into an overall usability evaluation.

### 4.1 Individual Usability Evaluations

Based on the usability categories defined by Nielsen [Nielsen-1994 (1)], the individual usability evaluations include efficiency, errors, and learnability. In each evaluation, a window in a sequence of extracted window switching is recognized as a step in a performed task. Every step is evaluated and assigned a score. In the evaluations, when the score is high, the step has usability problems. Below, we describe each usability evaluation considering a three-step task as an example.

#### 4.1.1 Efficiency Evaluation

Considering operation time to achieve a task, end-users usually require more time than expert users. NEM (Novice Expert ratio Method) is one method that focuses on this difference [Kurosu-2002]. NEM calculates the ratio of end-users' operation time to expert users' operation time. This ratio is called the "NE (Novice Expert) ratio". When the NE ratio is greater than 4.5, the GUIs of the task have problems.

In our method, the operation time is defined as the time each window is operated. The operation time for a window is from activation of the window to activation of the next window in window switching. The end-user operation time for a task is the average time of all the end-users. Figure 2 shows an example of the NE ratio. In this case, the NE ratio of window B exceeds 4.5. Consequently, window B is considered to have problems. It is noteworthy that window B also has the most end-users with an individual NE ratio above 4.5.

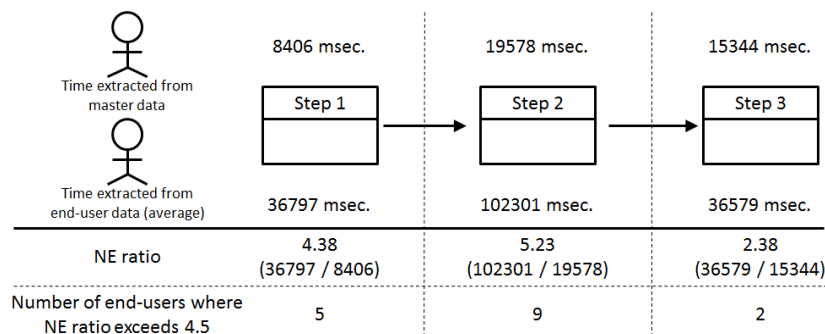


Figure 2. Example of the NE ratio

#### 4.1.2 Error Evaluation

Software where end-users often make mistakes is considered to have usability problems. Our method focuses on differences between operations of a task in the master and end-user data by comparing the two. In this evaluation, information of user events occurring in widgets and window switching is used. If the operations recorded in the end-user data differ from those recorded in the master data, then the score is added. In this evaluation, the sum of scores for each step of the task is calculated, and a high score indicates usability problems. Table 1 shows the score of each operation.



Table 1. Scores defined for error evaluation

Target operations		Score
Unintentional window switching		5
Unnecessary widget operation	With window switching	3
	Without window switching	1

Using the scores defined in Table 1, the sum of the scores for each step in a task is calculated from the end-user data. Table 2 shows an example of the calculated scores for each step. In this example, the sum of the scores for the step 1 is higher than the other steps.

Table 2. Example of calculated scores for each step

Target operations		Step 1	Step 2	Step 3
Unintentional window switching		6	2	2
Unnecessary widget operation	With window switching	3	0	2
	Without window switching	57	34	23
Sum of scores for each step		66	34	27

### 4.1.3 Learnability Evaluation

The first time end-users operate the software, it takes a long time, but once they become proficient, the operation time becomes shorter until reaching a constant level. If end-users have to operate the software many times to become proficient, then the software has usability problems. In our method, the operation time of each step in a task extracted from the master data recorded in Section 3.2 is considered to be the operation time that end-users become proficient. Thus, when end-users perform the same operation twice, the step time of the two operations are compared. In cases where the step time of the second operation is not close to that of the master data for the first operation, the step is deemed to have usability problems.

In this evaluation, following formula is used to evaluate learnability.  $T_e$  indicates the operation time of a step recorded in the master data,  $T_{u1}$  is the average time of the step in the first operation by all end-users, and  $T_{u2}$  denotes the average time of the step in the second operation by all end-users. If this calculated value is high, then the step has usability problems.

$$\frac{T_{u2} - T_e}{T_{u1} - T_e}$$

Table 3 shows an example of a learnability evaluation based on master and end-user data. In this example, the calculated value of step 1 is higher than the other steps.

Table 3. Example of calculated values for learnability

	Step 1	Step 2	Step 3
Operation time in master data (msec)	8406	19578	15344
Average time of first operation by end-users (msec)	30797	77301	36579
Average time of second operation by end-users (msec)	15484	30953	20375
Calculated values	0.316	0.197	0.237

## 4.2 Integration of Usability Evaluations

After performing three types of usability evaluations, the results are integrated to extract the more serious usability problems using AHP (Analytic Hierarchy Process) [Saaty-1980]. AHP is used to determine the optimal resolution by evaluating resolutions for a specific purpose. First, the evaluation criteria and their priorities are established. Then these are used to assess the resolutions. Next, the numerical values are calculated. If the value is high, then the resolution is considered to be optimal.

Our method aims to elucidate the more serious usability problems. Thus, the steps of a task are considered to be resolutions of AHP. Three usability categories, efficiency, errors, and learnability, are considered as criteria of AHP because their importance differs according to software developers, end-users and software characteristics. Figure 3 shows an example of a hierarchical structure of the relationship between usability categories and steps of a task along with AHP.

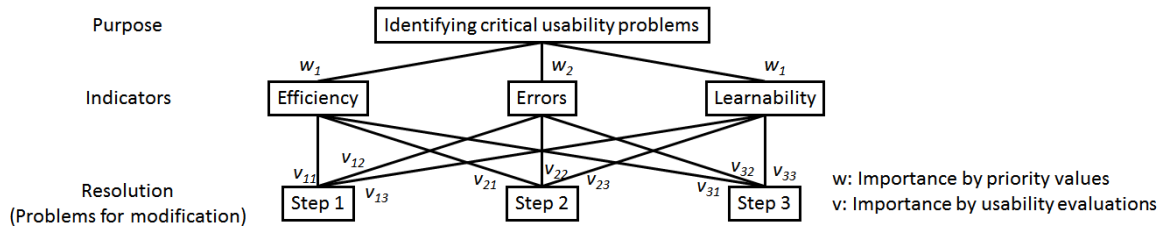


Figure 3. Example of hierarchical structure

In Figure 3, "w" is a value calculated using a pairwise comparison matrix, which is derived using priority values for criteria determined by software developers and end-users. Additionally, "v" is a value calculated using a pairwise comparison matrix, which is created using the scores of individual usability evaluations. For Figure 3, the pairwise matrix and the formula for calculating "w" are followings. In these, the priority values for the criteria "efficiency", "errors" and "learnability" are  $s_1$ ,  $s_2$  and  $s_3$ , respectively. Then,  $a_{ij} = s_i / s_j$ .

$$\begin{pmatrix} 1 & a_{12} & a_{13} \\ a_{21} & 1 & a_{23} \\ a_{31} & a_{32} & 1 \end{pmatrix} \quad w_x = \frac{\sqrt[3]{\prod_{j=1}^3 a_{xj}}}{\sum_{i=1}^n \sqrt[3]{\prod_{j=1}^n a_{ij}}}$$

Also, the pairwise matrix and the formula for calculating "v" are followings. In these, the scores of efficiency, errors and learnability evaluations for each step are  $r_{m1}$  (for step 1),  $r_{m2}$  (for step2) and  $r_{m3}$  (for step 3). "m" indicate 1 (efficiency), 2 (errors) or 3 (learnability). For example,  $r_{m1}$  indicates the result value of efficiency evaluation for step 1. Then,  $b_{mij} = r_{mi} / r_{mj}$ .

$$\begin{pmatrix} 1 & b_{m12} & b_{m13} \\ b_{m21} & 1 & b_{m23} \\ b_{m31} & b_{m32} & 1 \end{pmatrix} \quad v_{mx} = \frac{\sqrt[3]{\prod_{j=1}^3 b_{mxj}}}{\sum_{i=1}^n \sqrt[3]{\prod_{j=1}^n b_{mij}}}$$

Due to the individual usability evaluations, the sum of score for each step in a target task is calculated. However, the basis of the scores in each usability evaluation differs. Thus, these scores are modified as deviation scores, and Table 4 shows an example of the priority values for criteria and values of "w". In this Table 4, the most focused usability category is errors. The right of Table 4 is the pairwise matrix for calculating "w". Table 5 shows an example of values of "v".

Table 4. Example of priority values of criteria and "w"

	Efficiency	Errors	Learnability
Priority	2	7	5
Value of "w"	0.154 ( $w_1$ )	0.538 ( $w_2$ )	0.308 ( $w_3$ )

$$\begin{pmatrix} 1 & 2/7 & 2/5 \\ 7/2 & 1 & 7/5 \\ 5/2 & 5/7 & 1 \end{pmatrix}$$

Table 5. Example of values of "v"

	Step 1	Step 2	Step 3
Result of efficiency evaluation	0.327 ( $v_{11}$ )	0.403 ( $v_{21}$ )	0.270 ( $v_{31}$ )
Result of error evaluation	0.410 ( $v_{12}$ )	0.400 ( $v_{22}$ )	0.290 ( $v_{32}$ )
Result of learnability evaluation	0.406 ( $v_{13}$ )	0.275 ( $v_{23}$ )	0.319 ( $v_{33}$ )

Using Tables 4 and 5, the final scores of AHP are calculated using the following formula on the left of Table 6. If this formula yields a high score, the step has more serious usability problems compared to other steps. Table 6 shows the scores calculated using this formula. According to this table, step 1 has the highest score focusing errors. Consequently, step 1 is deemed to have the most serious usability problems.

$$p_i = \sum_{x=1}^n w_x v_{xi}$$

Table 6. Example of the final result

	Step 1	Step 2	Step 3
Integrated result	0.396	0.308	0.296

## 5. EVALUATION

To confirm effectiveness of our method, we evaluated the usability of two software packages: JWorkSheet [JWorkSheet] and task management system. The JWorkSheet is used to manage work and measure the time within a project, while the task management system is used to manage work and deadlines in a project. This evaluation was performed by 15 participants, who have high information literacy but are not expert users for these two software packages. The task for this usability evaluation was as follows:

- JWorkSheet
  - Step 1-1** Select a project and task previously registered, and begin timing measurement
  - Step 1-2** Register a new project and a task
  - Step 1-3** Select the project and task registered in Step 1-2, and output a report.
- Task management system
  - Step 2-1** Login using ID and password
  - Step 2-2** Add a task
  - Step 2-3** Register completion of a task
  - Step 2-4** Register deferment of a task
  - Step 2-5** Change the name of a task
  - Step 2-6** Print a list of all tasks

Table 7 shows the results of performing the above two tasks, while Table 8 shows the sums of scores calculated in individual usability evaluations. The parentheses in Table 8 indicate the values calculated using a pairwise comparison matrix based on the scores on the left.

Table 7. Results of performing tasks

	Operation time (msec)			Unintentional window switching	Unnecessary widget operations	
	Master data	First operation by end-users (average)	Second operation by end-users (average)		With window switching	Without window switching
Step 1-1	8406	30797	15484	6	3	57
Step 1-2	19578	77301	30953	2	0	34
Step 1-3	15344	36579	620375	2	2	23
Step 2-1	5578	20375	15407	1	1	0
Step 2-2	11735	36766	15781	9	3	14
Step 2-3	18500	54888	24172	2	0	9
Step 2-4	11625	54406	17125	2	1	11
Step 2-5	17109	37613	31563	2	2	13
Step 2-6	3641	16953	6969	4	3	7

Table 8. Results of individual usability evaluations

	Score of efficiency	Score of errors	Score of learnability
Step 1-1	5 (0.327)	96 (0.41)	0.316 (0.406)
Step 1-2	9 (0.403)	44 (0.3)	0.197 (0.275)
Step 1-3	2 (0.27)	39 (0.29)	0.237 (0.319)
Step 2-1	6 (0.324)	8 (0.263)	0.664 (0.414)
Step 2-2	4 (0.277)	68 (0.443)	0.162 (0.28)
Step 2-3	5 (0.342)	19 (0.335)	0.156 (0.323)
Step 2-4	12 (0.342)	21 (0.335)	0.129 (0.323)
Step 2-5	3 (0.26)	29 (0.324)	0.705 (0.416)
Step 2-6	9 (0.366)	36 (0.337)	0.25 (0.297)

In Tables 7 and 8, the results of individual usability evaluations were integrated based on priority values of the criteria. The priority values for criteria and "w" in Section 4.2 are shown in Tables 9, 10, and 11.

Table 9. Priorities of criteria (Focused on efficiency)

	Efficiency	Errors	Learnability
Priority values	9	1	5
Value of "w"	0.6	0.067	0.333

Table 10. Priorities of criteria (Focused on errors)

	Efficiency	Errors	Learnability
Priority values	5	9	1
Value of "w"	0.333	0.6	0.067

Table 11. Priorities values of criteria (Focused on learnability)

	Efficiency	Errors	Learnability
Priority values	1	5	9
Value of "w"	0.067	0.333	0.6

Based on Tables 9, 10, and 11, the individual usability evaluation results were integrated using AHP. Figures 4 and 5 show the integrated results for JWorkSheet and the task management system, respectively.

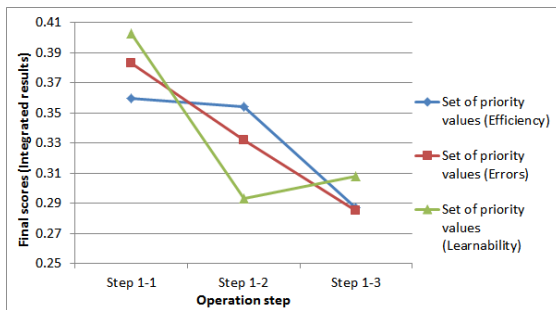


Figure 4. Integrated result of JWorkSheet

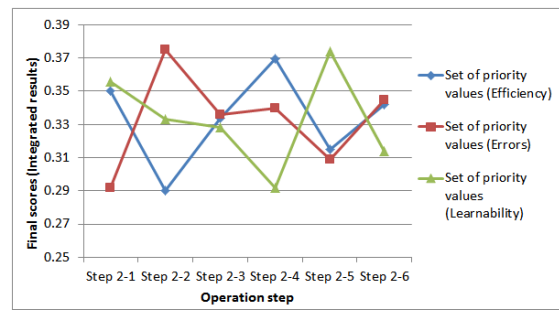


Figure 5. Integrated result of the task management system

These results imply that the focused steps with usability problems depend on the priorities of the criteria. For example, Figure 4 for JWorkSheet shows that step 1-1 had usability problems for errors and learnability because participants were unfamiliar with the time and work management system. However, even if we explained JWordkSheet prior to the evaluations, many participants seemed unable to concretely recognize their ability. Figure 4 also shows that step 1-2 had usability problems. Step 1-2 required the same type of data to be inputted, but operation flow had some limitations. Thus, participants appeared to have difficulty understanding the operation flow. However, step 1-2 had a good value of learnability, which is likely because the participants spent too much time on the first operation.

Figure 5 depicts large differences in the focused steps based on the priorities of the criteria. In the figure, the error values decrease from step 2-2 to step 2-5. In this task management system, after step 2-2, almost all the same windows were displayed, but the a few widgets differed in the windows. Hence, the values decreased because the participants became familiar with the system from step 2-2 to step 2-5.

In our method, a step is operations in a window. Consequently, it may be difficult to identify detailed usability problems in a window. That is, in this task management system, even if errors decreased from step 2-2 to 2-5, detailed usability problems may be present in the windows. However, our method needs to be improved so that steps can be subdivided, and usability evaluations can be performed at a level of widgets.

Regardless, our method can integrate individual usability evaluations, and steps with usability problems can be extracted based on the priorities. Moreover, the integrated results are appropriate, implying that our method can effectively extract more serious usability problems based on criteria established by software developers and end-users.

## 6. RELATED WORKS

Babaian et al. proposed a method to use operation histories of end-users to design interfaces [Babaian-2006]. In every component, various types of information such as operation time and keystrokes are recorded. Then

usability assessments are performed in terms of efficiency of UI operations and work achievements. However, the results are not integrated based on criteria that software developers and end-users determine. In our method, the results of usability evaluations can be integrated based on specific criteria.

Fukuzumi et al. proposed a method to evaluate usability of a system via checklists [Fukuzumi-2009]. The criteria are clear in the checklists and the evaluation can reflect the evaluator's intent. However, evaluations of a large scale system are a heavy burden and highly skilled evaluators are necessary to appropriately assess usability. In contrast, usability evaluations in our method only require the evaluators to operate the software.

Fiora et al. proposed to evaluate usability of software using component information and operation histories of end-users [Fiora-2008]. GUIs are evaluated based on operating time, number of times of component usage and users' behaviors. Based on this method, a usability testing framework for mobile terminals is developed. However, usability evaluations in this method are performed for one screen. In our method, usability of many windows is evaluated, and more serious usability problems can be identified.

Atterer et al. proposed a method to evaluate usability of web pages [Atterer-2006]. In this method, operation histories, such as mouse and keyboard events in web pages, are recorded using JavaScript. Based on the recorded information, usability evaluations are performed by analyzing whether end-users finish a specific task. However, evaluations of various perspectives are not performed in this method.

## 7. CONCLUSION

In this paper, we proposed a method to integrate various types of usability evaluations. Using our method, three types of usability evaluations are performed individually, and the results are integrated. By integrating the results of individual usability evaluations, more serious usability problems can be extracted based on the priorities of usability categories. Additionally, we evaluated our method and the detailed usability problems could not be identified. However, we could confirm the effectiveness of our method.

Future projects will focus on:

- Defining more appropriate scores for each individual usability evaluation
- Extracting more detailed usability problems
- Proposing various types of integrating usability problems
- Developing another usability evaluation method

## REFERENCES

- [Atterer-2006] Atterer, R., Wnuk, M., Schmidt, A., 2006. Knowing the User's Every Move - User Activity Tracking for Website Usability Evaluation and Implicit Interaction, Proceedings of the 15th international conference on World Wide Web (WWW'06), Edinburgh, Scotland, pp.203-212.
- [Babaian-2006] Babaian, T., Lucas, W., Topi, H., 2006. Making Memories: Apply User Input Logs to Interface Design and Evaluation, Proceedings of CHI'06 extended abstracts on Human factors in computing systems, Quebec, Canada, pp. 496-501.
- [Barnum-2001] Barnum, C. M., 2001. Usability Testing and Research, Longman, London, England.
- [Fiora-2008] Fiora, T. W. A., Baker, S., Warren, I., Dobbie, G., 2008. Automated Usability Testing Framework, Proceedings of the ninth conference on Australasian user interface (AUI'08), New South Wales, Australia, pp.55-64.
- [Fukuzumi-2009] Fukuzumi, S., Ikegami, T. and Okada, H., 2009. Development of Quantitative Usability Evaluation Method, Proceedings of the 13th International Conference on Human-Computer Interaction (HCI International 2009), California, USA, pp. 252-258
- [JWorkSheet] JWorkSheet, <http://jworksheet.ponec.net/>
- [Kurosu-2002] Kurosu, M., Urokohara, H. and Sato, D., 2002. A new data collection method for usability testing - NEM: Novice Expert ratio Method, Proceedings of Usability Professionals' Association (UPA 2002), Florida, USA.
- [Miles-2005] Miles, R., 2005. Aspectj Cookbook, Oreilly & Associates Inc, California, USA.
- [Nielsen-1994 (1)] Nielsen, J., 1994. Usability Engineering, Morgan Kaufmann, Massachusetts, USA.
- [Nielsen-1994 (2)] Nielsen, J., 1994. Usability Inspection Methods. Proceedings of CHI'94 Conference companion on Human factors in computing systems. Massachusetts, USA. pp. 413-414.
- [Saaty-1980] Saaty, T. L., 1980. Analytic Hierarchy Process, McGraw-Hill, New York, USA.

# TABLETOP USER INTERFACE FOR NAVIGATION IN VIRTUAL ENVIRONMENTS

Tatu Harviainen, Hanna Koskinen and Jari Laarni  
*VTT Technical Research Centre of Finland*  
*P.O. Box 1000 02044 VTT, Finland*

## ABSTRACT

In this paper, we present a multi-touch tabletop user interface for navigation in virtual environments and describe its preliminary evaluation. We have constructed a set-up where a multi-touch sensitive tabletop display is combined with a three-wall large-scale display system. In our set-up, the multi-wall large-scale display provides an immersive 3-D rendering of the virtual environment while the separate tabletop display provides a 2-D map view of the same data along with a user interface for controlling both views. We have developed three variations of navigation techniques for the display system and have conducted a user study to preliminarily evaluate the overall usability of the set-up and to compare the three proposed navigation techniques. Qualitative and quantitative results of the study demonstrate the prospects and limitations of the navigation techniques and also provide input for further development of the multi-display systems. This work originates from the need to develop new tools for industrial process control, and we conclude this paper by discussing the implications of the user study for this application area.

## KEYWORDS

Multi-touch tabletop and large-scale display interaction, virtual reality, navigation, multi-display environments.

## 1. INTRODUCTION

In most process control tasks adequate spatial and situational awareness are critical for effective operation of the system. Currently, there is an increasing use of digital technology in control room (CR) environments which have caused a number of challenges, both from the technical as well as users' point of view. As a part of the digitalization process, the physical control panels and consoles used before are replaced by their digital counterparts i.e. a keyboard and mouse controlled displays on which a vast amount of process information can be graphically presented. In this digitalization process, CR operators' secondary task demands have increased (e.g. navigation), and the development and maintenance of situation awareness has become more difficult because of the growing complexity of the control system and the large amount and variety of user interface elements (e.g. displays) required for carrying out the control operations.

Nowadays the development of screen technologies has made it possible to equip CRs with interactive and large sized displays. For example, many of today's CR set-ups already make use of large-scale displays in the presentation of process information. New advanced visualization techniques have made it also possible to provide operators with a new kind of presentations and views to the controlled entity (e.g. 3D visualizations, simulations, virtual and augmented reality). These technologies could potentially support more immersive user experience in process control environments (e.g. looking deep into the process and operating the systems remotely). Lately, there has been a growing body of studies on multi-touch technology. Although, it is still not in use at operating plants, some of the more conceptually oriented studies have shown that multi-touch technology may be valuable also in these environments. By combining strengths of sophisticated interaction qualities and high-resolution graphical displays users are able to interact more efficiently and reliably with computer systems. Furthermore, the advances in display technologies can augment the artefacts of our physical and social environments, such as tables and walls, with digital technology. The future CR environment designs have been expected to take advantage of these technologies and leverage people's natural abilities to act on and manipulate real-world objects.

The user study presented in this paper is an early step in our work in-progress aiming to draw a concept for future CR environments. As a part of this process we have developed a multi-touch tabletop called Affordance table. The development of the tabletop display started from the scratch by studying what kind of gestures people would prefer and what kind of associations they had on them when operating the multi-touch based interface. Based on the results of the first study, a set of appropriate interaction techniques for the tabletop were developed and the tabletop was integrated to work with the large-scale display system that provides a 3D view to a virtual environment. As its current state, the main aim of the present study was to conduct a preliminary evaluation of the overall suitability of a tabletop display and virtual reality (VR) system combination for data visualization and navigation tasks in CR environments.

## 2. RELATED RESEARCH

Several navigation methods and input devices used for navigation in VR have been described [3]. However, many of the methods and devices described suffer from usability drawbacks that have a negative impact on ease of learning, operation speed or vigilance. Use of tabletop display as a navigation tool in combination with VR system has been studied in various settings. Ajaj et al. compared usability of two tabletop user interface (UI) navigation techniques with one large-scale display [1]. Compared navigation techniques were based on egocentric and world reference maps similar to two of the navigation techniques (A and B) presented in this paper. In their approach tabletop UI detected only single touch events and additional hardware, i.e. joystick was used in combination with the tabletop. Examples of work where combination of 2D projection on a table with tangible objects has been used to control 3D visualization on a separate display have been described, e.g., in [5]. A general problem when using large-scale displays is to how to get an access to all display elements, that is, how to navigate on a display and select items. There are several solutions to this problem: Navigation can be improved by providing a direct access to all screen locations, e.g., with a help of a laser pointer or of a gaze-based input device [12]. Novel UI elements such as, e.g., Vision Wand, Vacuum, and Frisbee have been developed (for a review, see [8]). Input methods that are based on visual hand tracking have also been designed [9].

One of the most promising features of the multi-touch display technology is its ability to provide hands-on computing experience for the user of the system. Since we are skilful in direct manipulation of objects by using actively our both hands, it seems reasonable to try to design tools and interfaces relying on these existing skills. Many studies that have investigated the use of two-handed interaction techniques with computers have shown their potential to improve both the directness and the degree of manipulation as compared to the traditional Graphical User Interfaces (GUIs) in the WIMP (Windows, Icons, Menu, Pointer) paradigm [4]. The foreseen advantages of multi-touch systems, such as encouraging two-handed interaction, allowing parallel input, making interface elements directly accessible, and affording collaborative use, are all the main drivers of the ongoing development. Especially, the utilization of hand gesture and multi-finger input for navigation and target acquisition are key features when interacting with these displays. One frequently mentioned challenge is the accuracy of the operation with touch based displays [2]. Many of today's interfaces require and make possible the selection of very small targets, but with touch-based interfaces target selection and fine grained operation is restricted by the noise of underlying technology and lower tracking resolution as well as the relatively large touch area of a fingertip itself. The precision of touch-based interaction is further limited by the lack of tactile feedback and the occlusion problems. Due to these drawbacks, paying special attention to the design of navigation and target acquisition techniques is required in order to reach hands-on user experience. Many techniques have been suggested for increasing the accuracy of gesture and touch-controlled interfaces [5,11,13].

## 3. METHOD

The main aim of the study was to conduct a preliminary evaluation of the overall suitability of a tabletop display and VR system combination for data visualization and navigation tasks in control room environments. Our second aim was to investigate how the use of multi-touch tabletop impacts the use experience and users' immersion into the virtual environment.

Three different navigation controls were compared in order to study how the variations in the interaction techniques affect the users' experience and attentional focus during task performance. In the test scenario the users had to search for a target item and identify it. Both quantitative and qualitative information of their performance was gathered. During the test, the performances of the participants were video-recorded, and all the interviews were audio recorded. Two questionnaires were also completed, one measuring participants' user experience on the three navigation methods, another one measuring their presence experiences [6].

### 3.1 Participants

Six volunteers participated in the study of which two were females and four males. Five of them were right-handed and one was left-handed. All the participants were either researchers or academic students. They had little or no experience on the use of touch screens, and none of them were an enthusiastic computer game player. Four of the participants were familiar with the building which 3D model was used as an environment for the search task whereas two of them had never visited the building before.

### 3.2 Technical Apparatus

User study was conducted on a prototype system which combines multi-touch display with an immersive three wall large-scale display system, as seen in Figure 1. The large-scale display consists of three rear-projected display walls, used in monoscopic viewing mode. The angle between display walls is 120 degrees, and as a whole display walls provide approximately 160 degrees' field of view when operated from behind the tabletop display. The user is able to operate the multi-touch sensitive tabletop display in a standing position. The multi-touch display is composed of an off-the-shelf display (36" diameter) with 1360 x 768 resolution and a firewire camera array placed on top of the display frame. Multi-touch sensing is based on a proprietary camera-based touch tracking software.

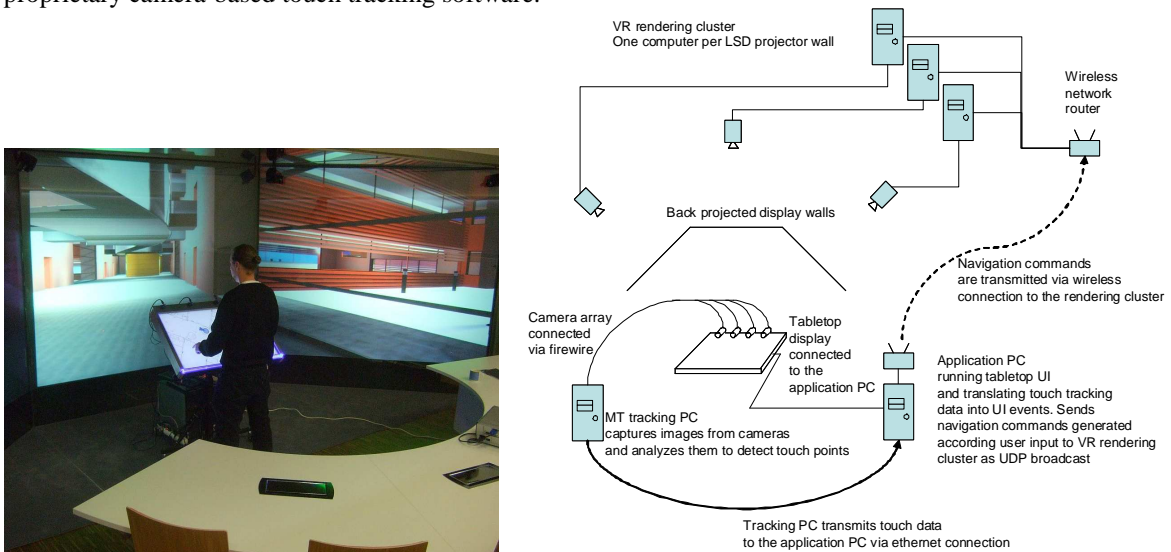
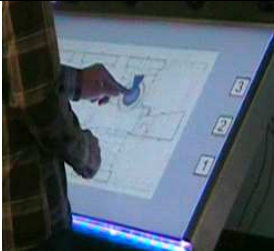
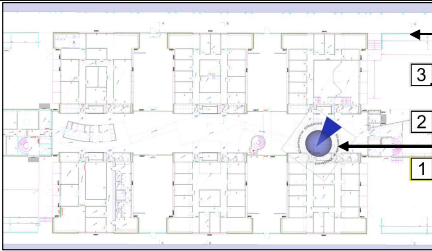

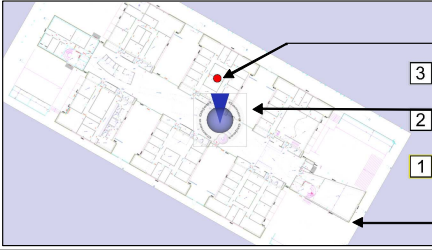

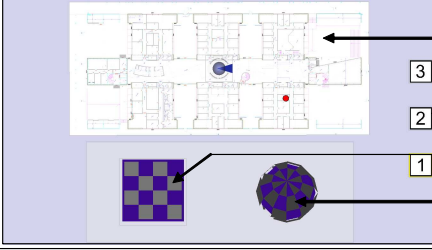


Figure 1 & 2. Overview picture of the system set-up and architecture

The total system is composed of tracking software for touch detection, a client application running on computer attached to the multi-touch display, rendering servers that display the virtual environment to the large-scale display and communication interfaces between the above-mentioned components as seen in Figure 2.



Table 1. User Interfaces of navigation techniques A, B and C.

<p><b>A</b></p> 	 <ul style="list-style-type: none"> <li>← 3 User controllable 2D map of the building</li> <li>← 2 Floor selection buttons</li> <li>← 1 User controllable camera object</li> </ul>	
<p><b>B</b></p> 	 <ul style="list-style-type: none"> <li>← 3 Navigation target visualized on the map</li> <li>← 2 Static camera object located on the centre of the screen</li> <li>← 1 Floor selection buttons</li> <li>← User controllable 2D map of the building</li> </ul>	
<p><b>C</b></p> 	 <ul style="list-style-type: none"> <li>← 3 2D map, camera location and target visualization</li> <li>← 2 Floor selection buttons</li> <li>← 1 Motion speed/direction control</li> <li>← View direction control</li> </ul>	

### 3.3 User Interface

The UI of the test set-up is distributed between tabletop display and large-scale display system. In all test cases, the large-scale display provides a panoramic 3D view to the virtual building model. In addition to the 3D view, the large-scale display does not include additional UI elements. The UI layout and elements on the tabletop were changed according to the tested interaction technique.

Three different navigation techniques were implemented for the test. In the following, the three different interaction techniques are labeled as A, B and C. Two of the implemented navigation techniques (A and B) were based on the map metaphor and used a 2D floor plan of the virtual building model as a part of the UI. The third navigation technique (C) was based on the vehicle control metaphor and the UI provided two interactive control elements. The basic controls for the three interaction techniques were designed on the basis of the results of an earlier test on multi-touch gesture preferences and associations on them [9].

#### 3.3.1 Navigation Technique A

In the navigation technique A, the user can inspect a 2D floor plan displayed on the tabletop display and navigate through it by manipulating a graphical camera object, which represents a camera location relative to the 2D floor plan (Table1). The floor plan can be rotated, scaled and translated freely with multi-touch gestures. The camera object controls the viewpoint of the 3D rendering, so that the user can freely navigate the 3D model of the building by manipulating it. The camera object can be translated with a single touch manipulation and rotated with a multi-touch gesture where the user first selects the camera object by touching it and then defines the rotation angle with a second touch anywhere on the display area.

#### 3.3.2 Navigation Technique B

In the navigation technique B, the user can translate, rotate and scale the 2D floor plan displayed on the tabletop UI (Table1). The camera object is static and in a fixed position and orientation in the centre of the screen, and navigation is controlled by translating and rotating the 2D floor plan. 2D floor plan is translated with a single touch manipulation and rotation and scaling can be activated simultaneously to the translation

with a multi-touch gesture. As the floor plan is translated or rotated, the viewing position and orientation of the virtual environment on the large-scale displays is updated accordingly..

### 3.3.3 Navigation Technique C

The navigation technique C is based on a control metaphor often used in computer games. This third navigation technique was included in the comparison because people tend to have more experience on using such controls, i.e. playing games and using radio controlled vehicles, than on using the techniques A or B. In this navigation technique two control elements are provided, one for manipulating the view orientation in the virtual environment and other for defining direction and speed of the translation of the view in the virtual environment. Controls can be thought of as being a steering control and a thrust direction / speed control. Also in this navigation technique, a static 2D floor plan and a static camera object are being visualized to provide visual information to the user about the view location in the virtual environment.

## 3.4 Procedure

### 3.4.1 Introduction and Training Phase

In the beginning of the test, participants were introduced with the main idea and purpose of the study followed by a short background interview. The training phase was divided into two sub-phases, training of the basic gestures and training of the navigation techniques. The first phase of the training focused on the use of multi-touch display and the basic multi-touch gestures. The goal of this phase was to introduce and evaluate the basic controls as well as capture the first reactions on the use of multi-touch tabletop. For this sub-phase the UI was modified so that each of the single multi-touch control was first used and trained individually (i.e. translation, rotation, and scale). More complex interactions allowing simultaneous combination of controls were introduced at the end of the training phase (Figure 3). The aim of this training approach was to allow users to learn the complete UI gradually but also to be able to evaluate each control type and associated gestures individually. During this phase the user's experience on the basic gestures were evaluated by a three-item questionnaire measuring the pleasantness, easiness and accuracy. In the second part of the training, the participants were introduced with the integrated VR displays and multi-touch tabletop system. They were also given a possibility to test and practice the use of the three compared interaction techniques (A, B and C) separately

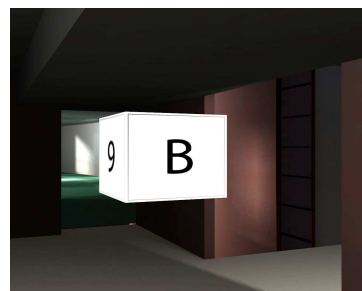
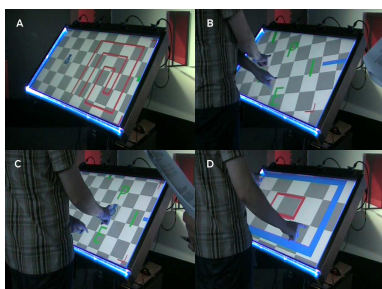


Figure 3. Training session on basic gestures. Figure 4. Search target visualized in immersive 3D environment.

### 3.4.2 Main Test Phase

In the main test phase, the participants carried out twelve search tasks. The test included two kinds of search and identification tasks. In one task the location of the cue was shown in the floor plan, and the users had to navigate to the given location of the building. In another task, no cue was given, and the users had to search for it within the building. Within-subjects design was applied in the test. The order of the trials was counterbalanced.

The basic task for the test was to navigate within the virtual building and to search cubes that were hidden around the building. After finding the cubes the participants were asked to identify the numbers and letters on the sides of the cubes to complete individual reach tasks (Figure 4). After each test run, the participants also completed the user experience questionnaire. The presence questionnaire was completed after six test runs.

### 3.4.3 Debriefing Phase

After the main test phase a debriefing interview was carried out to gather user's preferences on the integrated display system. The participants were also asked about the possible problems that they had encountered. The debriefing interview included questions around three different themes: 1) the potential usefulness of multi-touch technology; 2) user experience on the tested navigation techniques and the multi-touch display itself; and 3) the perception of the 3D virtual environment and the use of the integrated VR display and multi-touch tabletop system.

## 4. RESULTS

### 4.1 Quantitative Findings

The results of a one-way ANOVA showed that there was a significant difference between navigation methods in search time,  $F(2,33) = 4,1$ ,  $p < 0,05$ . According to post hoc tests, Method A was significantly faster than Method C. There was, however, no significant difference in search times between Method A and B or between Method B and C.

According to the questionnaire results, there was a significant overall effect of navigation method on user experience (all  $p_s < 0.01$ ). According to post hoc tests, Method A was considered easier and more pleasant and accurate than Method C. Method A was also thought to be in overall better suited for navigation than Method C. The participants also thought that Method B was easier than Method C: it was also considered in overall better than Method C. Interestingly, Method A was considered easier than Method B: it was also thought to be in overall better suited for navigation than Method B. In sum, these results suggest that the method in which the camera object is moved around relative to the map view is the best alternative of the three navigation methods that were tested.

According to the presence questionnaire, there was a significant overall effect of navigation methods on experienced action possibilities in the virtual environment,  $F(2,15) = p < 0.05$ . According to post hoc tests, the participants focused more intensively on the action possibilities within the 3D model of the building when using Method A than Method C. This finding suggests that when the camera itself is moved around by touch, the user has the best opportunity to focus his/her attention in the virtual environment.

### 4.2 Qualitative Findings

#### 4.2.1 User Experience on Basic Gestures, Multi-Touch Tabletop and the Navigation Methods

As mentioned above, the basic controls for the three interaction techniques were designed on the basis of the initial user study that were organized for assessing the preferred multi-touch gestures and associations on them. Generally, the participants agreed that the basic gestures that were used for the manipulation of the multi-touch display were intuitive and easy to learn, that is, the movements were immediately associated with the functions that they were designed for. From the basic gestures, the one serving the "moving the floor plan" function was preferred most. Also, the gesture for zooming functioned as the users expected. The participants had problems most often when they were asked to perform accurate operations (e.g. zooming objects to some given size).

According to debriefing interviews, method A was thought to be best suited for the navigational search task at hand. It was thought to be the most intuitive, and it was also considered to be the easiest technique to navigate through the building. According to the debriefings, the use of the other two techniques (B and C) caused some confusion among the participants as they did not always work as expected. Overall, there were more difficulties experienced in navigation when using methods B and C. However, the participants thought that also methods B and C had some useful properties; for example one participant mentioned that method C seemed to be potential for "steering" purposes as the way of using the control elements and 3D view reminded the way how the steering wheel and view out from the car's windscreen is used. But because of the

lack of the tactile feedback, attention had to be paid on the use of control elements, and the participants were therefore not able to fully utilize the 3D view when using method C.

The participants were also asked to evaluate the potential usefulness of multi-touch tabletop and the three introduced navigation techniques. They thought that the multi-touch technology would be the most suitable for navigation and browsing purposes. The system was also considered useful for design and problem solving tasks (e.g. viewing larger pictures/drawings). Particularly the manipulation (rotating, moving, scaling) capabilities were thought to be promising for these purposes.

#### **4.2.2 Sources and Types of Operating Problems**

Video recordings from the main test phase were also used for analyzing the magnitude and type of operating errors (technical or human originating). The number of errors was the lowest with method A. Generally, rotating the camera around the cube (operating the close-up view) caused most problems. It was typical that while trying to change the viewing angle of the camera (with method A) the camera accidentally drifted out of the building. In most cases this failure did not have dramatic consequences for the operation, but it came evident that when using the method A a lot of room was needed for executing the rotating action. Another typical error with the method A was that the participants accidentally manipulated the floor plan instead of the camera object.

The main difficulty with the interaction method B, in which 2D floor plan view at the bottom was moved relative to the static camera object, was that while trying to move and turn the floor plan picture the participants accidentally also zoomed it (usually it was zoomed out so that the 2D floor plan shrank). This action hampered the manipulation of the 2D floor plan view. Furthermore, because of this unexpected zooming effect and the other difficulties with the parallel manipulations (moving and rotating) the participants reported that they had got lost out of the map for short periods of time.

The number of mistakes was the highest when using interaction method C. Errors related to coordination were typical for this method. For example, the camera was accidentally moved or rotated to the wrong direction causing confusion. One reason for these mistakes might be the conflicting orientation of the 2D floor plan view on the tabletop and the 3D view on the large-scale display. Furthermore, also the manipulation of the control elements caused problems. It seemed that the controls did not react in the way the participants expected. Especially, the control for turning the camera was shown to be slow and inaccurate.

#### **4.2.3 User Experience on the integrated 3D Large Screen View and Multi-Touch Display System**

The participants thought that since the navigation task was quite easy and they had not attended to the 3D environment very intensively, the navigation experience was not very immersive. Participants were also asked to evaluate the potential usefulness of the combined multi-touch tabletop and the wall display system. They thought that in an industrial process control context, the multi-touch tabletop and the large-scale displays could complement each other, and the combined interactive display system could potentially provide an overall medium for the navigation and information presentation. The kind of presentation format (virtual environment) and way of interacting (multi-touch) was thought to be useful in situations in which there is a need for monitoring places and objects that are not approachable and easily viewable. The combined display system was also considered to be useful for remote operations.

## **5. CONCLUSION AND FUTURE WORK**

Overall, our test suggests that a multi-touch screen can be used as a remote control device for controlling movements of an object in a virtual environment. It was also found that some interaction methods are better suited for navigation in a virtual world than others. More research is however needed on determining the appropriate methods for multi-touch screen based navigation. The results of the test provide design implications for the further development of each navigation technique.

We are aiming to develop an interaction concept and a prototype for more efficient and integrated operation of industrial process (Figure 5). We expect that interactive surfaces are central media for the future CR environments and that this kind of integrated display systems can be designed to support collaborative management of the process. The future CR environments equipped with the combined multi-touch tabletop

and wall mounted large-scale displays should help operators to manage different types of situations as well as help them to maintaining the global overview and awareness of the process. In these environments digital information would be distributed over the whole physical CR space and its objects. New interface technologies enabling more concrete representation of complex phenomena and based on existing human manipulation skills can be expected to promote better awareness and sense of control. These technologies may also support more efficient and effective navigation within the complex system.



Figure 5. Concept images of the future control room settings.

## REFERENCES

- Ajaj, R., Vernier, F., Jacquemin, C., 2009, Navigation Modes for Combined Table/Screen 3D Scene Rendering. *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*. Banff, Alberta, Canada, pp. 141-148.
- Benko, H., Wilson, A., and Baudisch, P., 2006. Precise Selection Techniques for Multi-Touch Screens. *In Proceedings of CHI 2006*. Montreal, Canada.
- Bowman, D., Kruijff, E., LaVoila, J., Poupyrev, I., 2005. *3D User Interfaces: Theory and Practice*. Addison-Wesley 2005.
- Buxton, W. and Myers, B., 1986. A Study in Two-Handed Input. *In Proceedings of CHI'86*. pp. 321-326.
- Fiebrink, R., Morris, D., and Morris, M., 2009. Dynamic Mapping of Physical Controls for Tabletop Groupware. *Proceedings of CHI'09*. Boston. USA.
- Hartmann, T., Wirth, W., Schramm, H., Klimmt, C., Vorderer, P., Gysbers, A. & Böcking, S. The Spatial Presence Experience Scale (SPES): A Short Self Report Measure for Diverse Media Settings. Manuscript.
- Khan, A., Fitzmaurice, G., Almeida, D., Burtnyk, N., Kurtenbach, G., 2004. A Remote Control Interface for Large Displays. *Proceedings of the 17th annual ACM symposium on User interface software and technology*. Santa Fe, NM, USA.
- O'Hara, K., Perry, M., Churchill, E., Russel, D. 2003. *Introduction to Public and Situated Displays In Public and Situated Displays*. Kluwer Publishers. Dordrecht, The Netherlands.
- Koskinen, H. Laarni, J., Honkamaa, P., 2008. Hands-on the Process Control: Users Preferences and Associations on Hand Movements. *Proceedings of CHI 2008*. Florence, Italy.
- Malik, S., Laszlo, J., 2004. Visual touchpad: a Two-Handed Gestural Input Device. *Proceedings of the 6th international conference on Multimodal interface*. State College, PA, USA.
- Nacenta, M. et al., 2009. Separability of Spatial Manipulations in Multi-touch Interfaces. *ACM, In Proceedings of Graphics Interface 2009*. Kelowna, British Columbia, Canada, pp. 175-182.
- Vogel, D., Balakrishnan, R., 2005. Distant Freehand Pointing and Clicking on Very Large, High Resolution Displays. *Proceedings of the 18th annual ACM symposium on User interface software and technology*. Seattle, WA, USA.
- Wu, M. and Balakrishnan, R., 2003. Multi-Finger and Whole Hand Gestural Interaction Techniques for Multi-User Tabletop Displays. *Proceedings of UIST '03*. pp.193-202.

# EXPOSING REPETITIVE SCANNING IN EYE MOVEMENT SEQUENCES WITH T-PATTERN DETECTION

Marcus Mast and Michael Burmester

*User Experience Research Group*

*Stuttgart Media University*

*Wolframstr. 32, Stuttgart, Germany*

## ABSTRACT

When analyzing eye movement it can be of interest to identify common ground in gaze sequences, for example to find cross-person similarities in the way people scan or to investigate scanning habits of a single person repeatedly exposed to a stimulus. Measuring and exposing such similarities is challenging from a methodological perspective, as the spatial and temporal dimensions have to be considered simultaneously. Techniques like string-editing can measure the similarity of gaze sequences but techniques for exposing the actual similarity are not yet established. In this paper, we propose the T-pattern technique (Magnusson, 2000) for the analysis of spatiotemporal eye movement data and describe its application. The technique has the ability to extract recurrent patterns of visual scanning from real-time region-coded fixation sequences. Extracted patterns can be qualitatively examined, allowing to derive hypotheses on the underlying cognitive processes of viewers. The technique also provides metrics for the repetitiveness of visual scanning, allowing for quantitative comparisons between experimental groups.

## KEYWORDS

Eye movement, eye tracking, scanpath, sequence analysis, T-pattern

## 1. INTRODUCTION

The analysis of eye movement can provide valuable data for understanding users' information processing and behavior. It can reveal how users search, read, or where they direct attention in a graphical user interface (e.g., Stark Adam et al., 2007). Most commonly, studies employ metrics and visualizations that collapse data over time, such as the mean fixation duration, transition frequencies, or attention maps. The process of scanning is studied less often (some examples being Josephson & Holmes, 2002; Pan et al., 2004; Grier et al., 2007). This may in part be due to the challenges involved in analyzing sequences of eye movement. Temporal and spatial aspects must be considered simultaneously for multiple gaze sequences. Figure 1 shows the visual scans of three people looking at a web page for 10 seconds each. While it seems that some similarities exist between the sequences (e.g., in the elements scanned or in local subsequences), identifying and judging concrete similarities is difficult for a human observer. The problem is aggravated with a growing number of sequences to be compared or when searching in long sequences. Sophisticated methodology is required for identifying and measuring similarity in eye movement sequences.

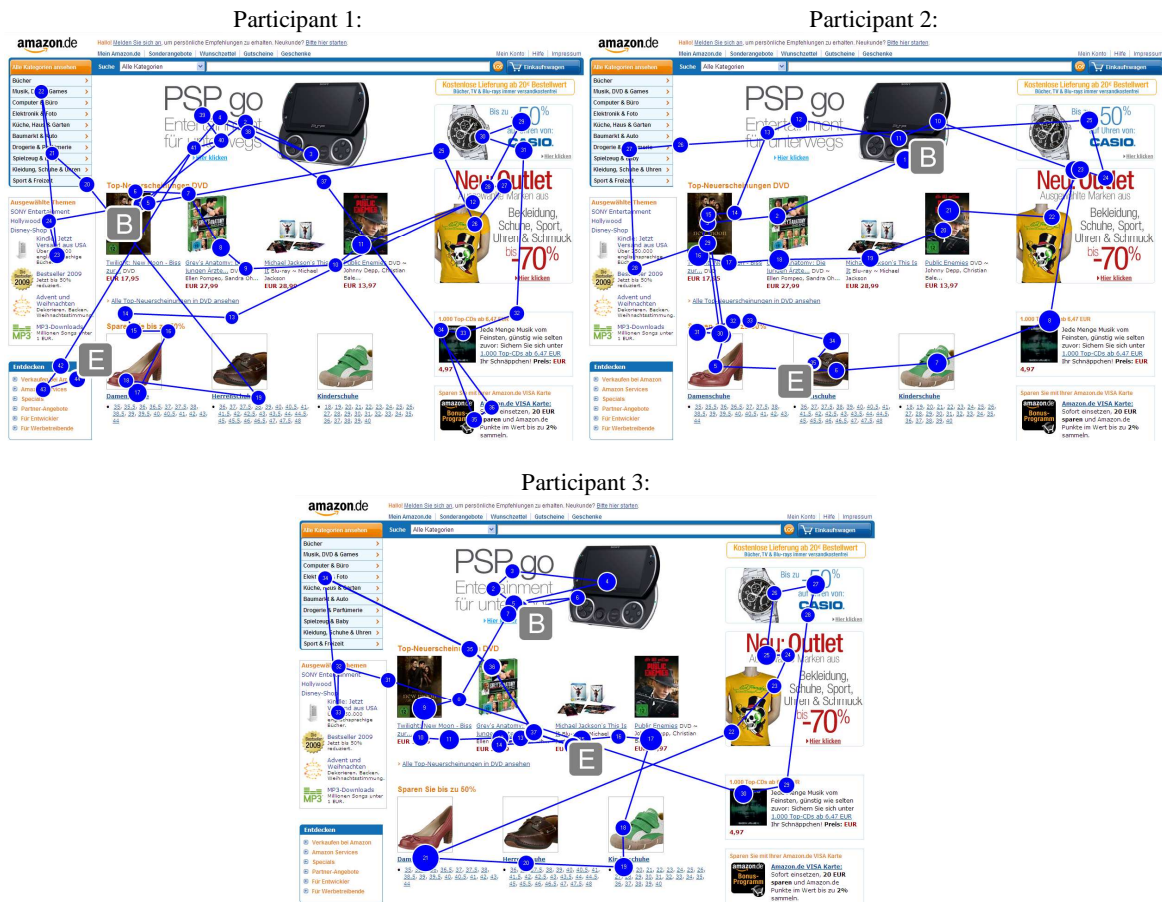


Figure 1. Three 10-second gaze sequences on the home page of online retailer Amazon (www.amazon.de); B denotes beginning and E end of a gaze sequence

## 2. STATE OF THE ART

Our primary interest is to expose concrete sequence similarities in a set of eye movement sequences (qualitative perspective). Secondly, we are interested in measuring similarities (quantitative perspective). In this section, we review existing techniques and techniques under development for these two aspects.

Perhaps the largest body of literature dedicated to eye movement sequences is concerned with Noton and Stark's (1971) scanpath theory. In this literature, Markov analysis (e.g., Stark & Ellis, 1981; Pieters et al., 1999) and string-editing (e.g., Brandt & Stark, 1997; Josephson & Holmes, 2002; Pan et al., 2004, Foulsham & Underwood, 2008) were mostly applied. Markov analysis (Markov, 1906/1971, Ross, 1996) does not compare actual eye movement sequences but tests the fit of a hypothesized region transition matrix against an observed one. It can neither qualitatively expose sequence similarities nor quantify them and is therefore unsuitable for either of the described purposes. The string-edit technique (Levenshtein, 1965/1966, Sankoff & Kruskal, 1983) measures the similarity of two sequences by transforming one into the other. Fixation sequences can be represented by a letter for each fixated region of interest (ROI). String-editing requires an order-only time scale, not taking into account real-time. The number of editing operations (deletions, insertions, substitutions) required to transform one sequence (or string) into the other serves as the measure of similarity. While groups of similar sequences can be identified by applying multidimensional scaling (Josephson & Holmes, 2002), string-editing does not expose actual similarities. Thus, string-editing can satisfy the quantitative requirement but not the qualitative. Some weaknesses of the metric are discussed in Section 4.3. There are further techniques only capable of measuring, not of exposing similarities (e.g., Gbadamosi & Zangemeister, 2001; Jarodzka et al., 2010).

Early work on methods that allow for a qualitative inspection of gaze sequence similarities was shown by Hembrooke et al. (2006). Their multiple-sequence averaging technique calculates a common gaze sequence from several input sequences. No detailed information on its procedure was provided and the technique remained to be validated. It also seems questionable if a single average path should be calculated from widely heterogeneous gaze sequences.

West et al. (2006) implemented three approaches for finding qualitative similarities in their "eyePatterns" software. The first, based on the work of Smith and Waterman (1985), aligns two sequences according to the position where their similarity is highest. This is suitable for finding the substring of highest similarity but due to the chronological form of the alignment it fails to find more than that. For example, in the two sequences ABCDVWXYZ and VWXYZABCD, only the longest common string VWXYZ is found but not ABCD. Another limitation is that the technique can only compare two sequences at a time. This makes it difficult to find patterns shared by multiple sequences. The second approach discovers common strings in multiple sequences but currently has no tolerance for substring deviation implemented, so, for instance, in ABCYDEF and ABCZDEF, the common pattern ABC-DEF cannot be found. The third approach incorporates gap tolerance but it only finds specifically entered strings and not unknown ones. All three approaches require a reduction of the temporal dimension to order-only. Thus, fixation durations or the length of dwelling in regions are not considered. Overall, the set of techniques shows potential but also plenty of room for improvement with regard to exposing similarities.

Goldberg and Helfman (2010) showed work on a new technique for finding clusters of similar gaze sequences. It requires order-only input data but it is not limited to pairwise comparisons. The technique first creates dotplots of concatenated gaze sequences. In the dotplots it searches for statistically significant ROI sequence matches between all possible combinations of sequence pairs through linear regression. Hierarchical clusters are then formed from qualitatively similar regressions. These can be examined in an interactive dendrogram. A current limitation is that one gaze sequence can only be assigned to one cluster, so two different scan patterns within a person's gaze sequence cannot be identified, even if similar patterns existed in other sequences. As for the quantitative aspect, Goldberg and Helfman's current focus does not seem to be on similarity metrics but it would be conceivable to form metrics, e.g. based on the amount or size of clusters. Despite some current limitations and the non-availability of the technique, Goldberg and Helfman's (2010) approach seems promising as it has the ability to expose sequence similarities across large numbers of eye movement sequences, incorporating tolerance for sequence divergence.

We subsequently describe the application of the T-pattern technique (Magnusson, 1983, 2000, 2005, 2006) to the analysis of spatiotemporal eye movement data, which has been used by Unz et al. (2005) and Burmester and Mast (2010). The technique extracts recurrent patterns based on invariant distribution characteristics of sequence elements. It uses real-time input data, incorporates tolerance to divergence between sequences, and has no limitations regarding the number of sequences to be compared or the number of qualitative similarities to be extracted.

### 3. T-PATTERN DEFINITION AND DETECTION PRINCIPLE

The T-pattern approach originates from the behavioral sciences and was originally developed for the analysis of social interaction. It has already been applied in a variety of areas such as ball and player movement analysis in soccer matches (Borrie et al., 2002) or DNA structure analysis (Magnusson, 2002). The assumption of the method is that behavior has a tendency to organize in a recurrent manner. Elicitation of unknown, non-obvious behavioral patterns can provide new knowledge on behavioral systems. The term T-pattern stands for temporal pattern. The detection algorithm has been developed by Magnus S. Magnusson during the past decades (Magnusson, 1983, 2000, 2005, 2006) and is commercially available in the Theme software (Magnusson, 2004).

T-pattern detection assumes input data coded on a one-dimensional discrete scale. Typically, this would be a real-time temporal scale but the scale can also be order-only or based on geodesic distance. The detection algorithm is scale-independent, i.e. the analyzed events may span centuries or occur within a microsecond. The input data for T-pattern detection needs to be coded as events. Following Magnusson (2000), a type of behavior will subsequently be referred to as an "event type" and an instance (an occurrence) of behavior will be referred to as an "event". An example for an event type would be "passes ball" and an



associated event would be “passes ball at 20m:21s”. An event only happens once, at a particular point in time and has no duration. Basic to the definition of a T-pattern is the invariance of the temporal distribution of events types. Two event types are considered a T-pattern if (1) they both occur at least twice in the behavior record in the same order and (2) their occurrences are invariantly distributed over time. Whether or not a distribution is considered invariant is determined through two possible types of so-called critical interval (CI) relationships. A *fast* critical interval is present if event type A is followed relatively quickly by event type B (Figure 2, top). A *free* critical interval does not require event type B to follow quickly but the time distance between A and B must be similar (Figure 2, bottom).

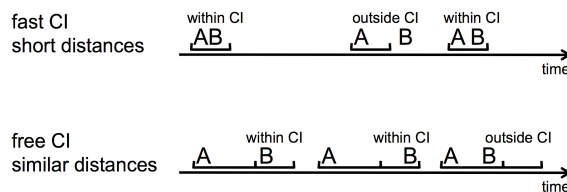


Figure 2. Fast and free critical interval (CI) relationships between event types

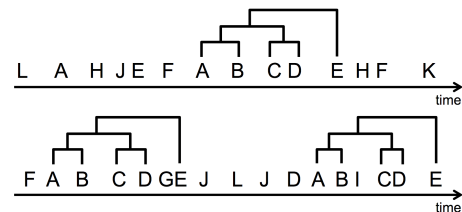


Figure 3. A T-pattern (ABCDE) occurring three times in two behavior records

In an iterative search process, each detected T-pattern can be complemented by another event type or another T-pattern to form a longer, higher-order T-pattern. A T-pattern with  $m$  components  $X_1..X_m$  can be formally expressed as a recurring ordered set:  $X_1 [d_1, d_2], X_2 [d_1, d_2], .. X_i [d_1, d_2], X_{i+1} .. X_{m-1} [d_1, d_2], X_m$  where  $[d_1, d_2]$  is the critical interval within which one component is followed by the next;  $d_1$  is the temporal distance from the component after which the interval starts and  $d_2$  the end of the interval (Magnusson, 2005). When searching for fast critical intervals,  $d_1$  is set to zero.

The relationship between two pattern components can thus be described as  $X_{left} [d_1, d_2] X_{right}$ , i.e. the left component  $X_{left}$  is followed within the critical interval  $[d_1, d_2]$  by the right component  $X_{right}$  (Magnusson, 2005). The critical interval algorithm uses the following procedure: “The critical interval algorithm measures the time from each occurrence (end) of  $X_{left}$  to the first following or concurrent occurrence (beginning) of  $X_{right}$ . Using this distribution and some preset significance level, it searches for the longest possible interval  $[d_1, d_2]$  such that  $(X_{left})$  (ending at  $t$ ) is, significantly more often than expected by  $h_0$ , followed within  $[t+d_1, t+d_2]$  by the beginning of another component ( $X_{right}$ ); where  $h_0$  is that  $(X_{right})$  is independently randomly distributed over the observation period  $[t_1, t_2]$  with a constant probability per time unit =  $N(X_{right}) / (t_2-t_1+1)$ ; where  $N(X_{right})$  is the number occurrences of  $X_{right}$ .” (Magnusson, 2005, p. 11).

The extent to which event distances may vary to still be recognized as a T-pattern is thus specified by the significance level ( $\alpha$ ). Few and shorter patterns are detected with a strict significance level while more and longer patterns are detected with a loose significance level. Figure 3 shows a T-pattern of length 5 (ABCDE) occurring three times in two behavior records. The T-pattern consists of the sub-patterns (AB) and (CD) which are connected as ((AB)(CD)) one level higher and as (((AB)(CD))E) on the highest level. The levels represent the algorithm iterations. While T-pattern elements always occur in the same order in the data and within the critical interval, other events may occur intermediately such as event type G occurring between D and E in the first pattern occurrence in the lower behavior record of Figure 3. In the iterative search process, only the longest patterns survive while shorter sub-patterns are discarded. The search process stops at the point where no more critical interval relationships can be found.

## 4. ANALYZING EYE MOVEMENT FOR T-PATTERNS

### 4.1 Data Transformation

Before eye movement data can be analyzed for T-patterns, it needs to be translated to event data. The starting point is region-coded fixation data as normally can be exported from contemporary eye tracker software, i.e. raw sampled x/y positional data has undergone a filter with a temporal and spatial threshold for eye immobility to be considered a fixation and a form of spatial data reduction, reducing the number of possible

fixation positions. Most typically, spatial data reduction would be achieved by defining ROIs but it can also be achieved by algorithmic segmentation or clustering (e.g., Privitera & Stark, 2000). While using a real-time scale retains most accuracy, this is not a requirement. An order-only scale can be used, e.g. if the research question demands disregarding region dwell times. There are several possible approaches for translating the region-coded fixation data: (1) All beginnings and all ends of fixations are retained, including repeated fixations on the same region; (2) Only *enter region* and *leave region* events (i.e., the beginning of the first fixation and the end of the last in a region) are retained; (3) Only *enter region* events are retained, i.e. all successive fixations within a region are collapsed into a single event occurring at the time of first fixation in the region. More translation alternatives are conceivable by using modifiers. Modifiers further break down event types. They can be used e.g. to differentiate fixation durations (e.g., *fixate,short*, *fixate,medium*, *fixate,long*) or to define ROI levels (e.g., *fixate\_rightcolumn,box1,picture*). Ultimately, the choice for an alternative depends on the research question. Approach 1 retains most accuracy while approach 3 offers a straightforward interpretation of detected T-patterns.

## 4.2 Detection Parameters and Pattern Filters

When searching for T-patterns in the transformed eye movement data, several considerations on detection parameters and pattern filters have to be made which have implications for the number, length, and content of detected T-patterns:

Types of critical intervals permitted: Fast only or both fast and free critical intervals (see Section 3) can be searched for. When allowing free CIs, pattern elements may lie farther apart. One may consider allowing free CIs if global scanning strategies over longer periods are of interest rather than just patterns with elements occurring in relatively short succession.

Significance level ( $\alpha$ ): The significance level specifies the maximum accepted probability of two event types considered a T-pattern to be within the critical interval by chance (see Section 3). A loose significance level (a higher value) will result in longer critical intervals and therefore more detected T-patterns. Since the detection process is iterative, in the next iteration, the loose level again will lead to more detections, extending the length (number of ROIs) of the previously detected T-patterns. Thus, at the end of the process more and longer T-patterns, consisting of more ROIs, will be detected with a loose significance level than with a stringent one. However, with a loose significance level the detected T-patterns will also have a higher likelihood to have occurred by chance. Because the significance level affects the length of CIs, when searching for fast CIs, it also affects how many ROIs a user may have looked at between T-pattern elements.

Minimum number of occurrences: To be considered a T-pattern, a pair of event types within the critical interval needs to occur at least twice. Specifying a higher minimum number of pattern occurrences results in a lower number of patterns found. When searching in multiple observation periods (for example in several stimulus viewings), it can be useful to additionally specify in how many independent observations each T-pattern must occur through the minimum samples parameter. For example, it can be specified that each T-pattern must occur in at least 50% of the stimulus viewings. Patterns with a high number of occurrences tend to be shorter than such occurring less often.

Minimum pattern length: By definition, each T-pattern must have a minimum length of two elements (ROIs). After detection, patterns can be filtered for higher minimum lengths (e.g., 3 or 4). Long patterns are usually found less often than short ones but may be more relevant, depending on the research question.

Specifying further parameters and applying further filters can be useful, e.g. to reduce redundancy or to qualitatively select T-patterns by behavioral content (e.g., a specific ROI contained).

## 4.3 Exemplary Qualitative and Quantitative Detection Results

To illustrate the capabilities of the T-pattern technique, we present exemplary results of a T-pattern search in the three eye movement sequences shown in Figure 1. The eye movement sequences represent the visual scans of three persons from a web usage study and are thus not fictive data. However, they were selected according to demonstration criteria. After defining ROIs and transforming the three sequences into event data using approach 3 described in Section 4.1, T-pattern detection was applied with  $\alpha = 0.01$ , fast CIs only, min. occurrences = 2, min. length = 2 ROIs. The chosen significance level offered a reasonable balance between chance occurrence avoidance and detection result relevance. The lowest minimum occurrence and length

values of 2 were chosen to allow an unrestricted view on all detected patterns. As we were not interested in patterns with long time distances between pattern elements, we did not search for patterns of the free type.

Overall, 11 T-patterns were detected in the three gaze sequences. There was 1 pattern of length 7 (2 occurrences), 1 of length 6 (2 occurrences), 1 of length 5 (3 occurrences), 2 of length 4 (2 occurrences each), 1 of length 3 (2 occurrences), and 5 of length 2 (between 2 and 4 occurrences). All three participants were included in 4 patterns. Participant 1 was included in 10 patterns, participant 2 in 8, and participant 3 in 8 too. A T-pattern can also occur repeatedly in a single sequence. For the present data, only short 2-ROI patterns occurred repeatedly within a participant's sequence (two 2-ROI patterns occurred twice for participant 3 and one for participant 2). Longer within-sequence T-patterns would more likely occur in eye movement records of longer duration, e.g. in an extensive search process of a user where elements are often re-scanned.

After detection, patterns can be inspected for included elements and occurrence time characteristics. Figure 4 shows the longest detected pattern. The pattern tree graph in the upper left shows the iterative detection structure of the T-pattern. The graphs in the upper right and bottom show the pattern's occurrences over time on the millisecond scale used for detection (participant 1: 0-9,999ms; participant 2: 10,000-19,999ms; participant 3: 20,000-29,999ms) with the upper right graph additionally showing sub-pattern occurrences. It can be observed that the pattern occurred once in the sequence of participant 1 (near the beginning), once in the sequence of participant 2 (in the middle), and it did not occur in the sequence of participant 3. Figure 5 visualizes the T-pattern over the stimulus and ROIs. Figures 6 and 7 visualize the T-pattern over the original fixation data. Figure 8 shows the pattern's occurrences in the event data. Two intermediate events (DVD\_heading, DVD2\_pic) occur for participant 1 and one (DVD1\_text) for participant 2.

The longest T-pattern occurring for all three participants (once for each) was 5 ROIs long. Figure 9 visualizes this pattern over the fixations of participant 3. Although it consists of the same first 5 ROIs as the pattern in Figure 4 and is based on the same event occurrences, it is formally not a direct sub-pattern because it has a different pattern tree structure (not depicted) due to the different ways a T-pattern can be iteratively constructed. When disregarding the detection tree, however, this pattern can be considered a subset of the longer 7-ROI pattern. The 5-ROI pattern in Figure 9 represents the largest common ground found across the three sequences. Most of the other nine T-patterns found are variations of the depicted ones, e.g. shorter versions with a different tree structure or including a different ROI and lacking another. Only one other pattern seems noteworthy. It occurred once each for participants 1 and 3 and has the following order of ROIs: text of the shoe in the middle, then box 3, box 2, box 4 in the right column. The high number of redundant and similar patterns is a common artifact of the T-pattern detection process and can obstruct a clear view on the relevant results, especially when working with larger datasets. Also, if behavior is very repetitious, many event type relationships within the critical interval can be found and in the iterative detection process, patterns can be formed in many different ways (with different pattern trees), leading to pattern redundancy. The implemented redundancy reduction mechanisms often do not sufficiently or adequately reduce patterns.

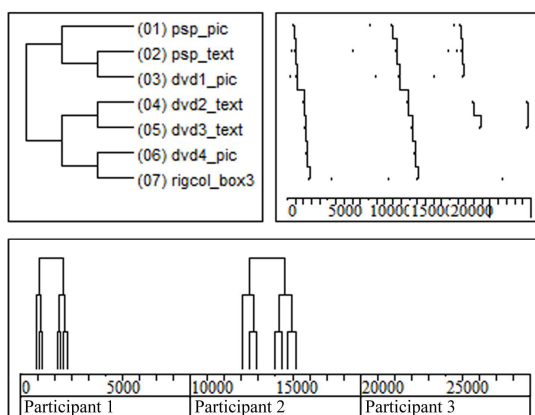


Figure 4. Longest T-pattern detected (length = 7 ROIs), occurring once each for participants 1 and 2

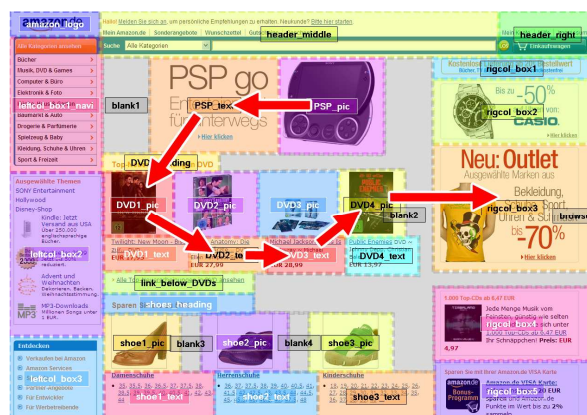


Figure 5. Visualization of T-pattern from Figure 4, ROIs overlaid on stimulus



## 4.4 Validation of Results

To validate the detected T-patterns against chance occurrence, a T-pattern search can be performed with the same search parameters in observation periods of equal length with the same fixation events but occurring in randomized order (Magnusson, 2004), i.e. a random path across the same fixation events. The result of this search can then be compared with the actual result. Validating longer patterns is particularly important as conclusions on behavior are usually derived from them. For the present exemplary data, results were compared to the average result of 1000 searches in randomized data. The searches in randomized data resulted in substantially less patterns found (7-ROI patterns: 1 real vs. 0.0 mean random, 6-ROI patterns: 1 real vs. 0.0 mean random, 5-ROI patterns: 1 real, 0.1 mean random, 4-ROI patterns: 2 real, 0.3 mean random), giving confidence in the validity of the results.

## 5. CONCLUSION

The main benefit of the T-pattern technique is its ability to expose recurrent visual scanning. The T-pattern technique is insensitive to noise elements in the data and can compare an unlimited number of sequences in one analysis run. Detected T-patterns may show well-known and expected behavior (e.g., top-to-bottom scanning of a column on a web page) or reveal unexpected or previously undiscovered recurrent scanning (e.g. inclusion of unexpected elements or an unusual scanning order). Previously unknown patterns allow forming hypotheses on the cognitive processes behind them. The metrics T-pattern length, occurrences, and quantity measure various aspects of repetition in eye movement sequences.

From the perspective of eye movement data, some enhancements of the T-pattern technique could be beneficial: (1) Sometimes a high number of near-redundant T-patterns detected can obstruct a clear view on the results. Redundancy reduction based on qualitative pattern similarity could help to further reduce patterns. For example, two patterns (XWZ) and (XYZ) could be consolidated as (X(W/Y)Z) and visualized as a branched structure. (2) The number of ROIs that may occur and the time interval that may pass between the ROIs of a pattern is determined indirectly through the significance level. The ability to additionally set hard upper limits would provide more certainty as to how far apart pattern elements may be in the data. (3) Spatial distance limits to intermediate ROI occurrences could be beneficial in some situations, e.g. to disregard T-patterns if intermediate ROI fixations outside a certain radius occurred. (4) Another functional enhancement would be more advanced filtering tools, e.g. to select patterns based on their occurrences in a specific sample of the dataset (e.g. a specific participant). Some improvements like pattern consolidation and the selection of patterns by samples are under consideration for future integration in the software (M. S. Magnusson, personal communication, December 2010) along with additional pattern detection approaches (Magnusson, 2006).

T-pattern detection has the ability to elucidate eye movement data in a number of situations. When persons repeatedly view a stimulus, T-pattern detection can help to understand what kind of visual scanning habits they form, if any (Burmester & Mast, 2010). Search processes can be analyzed for recurrent subsequences. Typical patterns of scanning can be compared between different experimental groups. For example, the effect of factors like task, stimulus design, user expertise, gender, or culture on the content of T-patterns and on their length, occurrences, and quantity can be investigated. Outside the domain of eye movement analysis, the technique can generally be considered in human-computer interaction research in cases where recurrent behavior is of interest, e.g. for the analysis of clickstreams, interaction in computer games, or computer-supported cooperative work.

## REFERENCES

- Borrie A., Jonsson, G.K., Magnusson, M.S., 2002. Temporal pattern analysis and its applicability in sport: an explanation and exemplar data. *Journal of Sports Sciences*, 20, 845-852.
- Brandt, S.A., Stark, L.W., 1997. Spontaneous eye movements during visual imagery reflect the content of the visual scene. *Journal of Cognitive Neuroscience*, 9, 27-38.

- Burmester, M., Mast, M., 2010. Repeated web page visits and the scanpath theory: a recurrent pattern detection approach. *Journal of Eye Movement Research*, 3(4):5, 1-20.
- Foulsham, T., Underwood, G., 2008. What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition. *Journal of Vision*, 8, 1-17.
- Gbadamosi, J., Zangemeister, W.H., 2001. Visual imagery in hemianopic patients. *Journal of Cognitive Neuroscience*, 13, 855-866.
- Goldberg, J.H., Helfman, J.I., 2010. Scanpath clustering and aggregation. *Proc. Eye Tracking Research and Applications (ETRA)*, pp. 227-234.
- Grier, R.A., Kortum, P., Miller, J.T., 2007. How users view web pages: an exploration of cognitive and perceptual mechanisms. In: Zaphiris, P., Kurniawan, S. (Eds.), *Human Computer Interaction Research in Web Design and Evaluation* (pp. 22-41). Hershey: Idea Group Publishing.
- Hembrooke, H., Feusner, M., Gay, G., 2006. Averaging scan patterns and what they can tell us. Late breaking results: oral presentations. *Proc. Eye Tracking Research and Applications (ETRA)*, p. 41.
- Jarodzka, H., Holmqvist, K., Nyström, M., 2010. A vector-based, multidimensional scanpath similarity measure. *Proc. Eye Tracking Research and Applications (ETRA)*, pp. 211-218.
- Josephson, S., Holmes, M.E., 2002. Attention to repeated images on the World-Wide Web: another look at scanpath theory. *Behavior Research Methods, Instruments, & Computers*, 34, 539-548.
- Levenshtein, V.I., 1966. Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*, 10, 707-710. (Original work published 1965)
- Magnusson, M.S., 1983. Theme and Syndrome: two programs for behavior research. In: Edwards, D., Hoeskuldsson, A. (Eds.), *Symposium in Applied Statistics* (pp. 17-42). Copenhagen: NEUCC, RECKU, & RECAU.
- Magnusson, M.S., 2000. Discovering hidden time patterns in behavior: T-patterns and their detection. *Behavior Research Methods, Instruments, & Computers*, 31, 93-110.
- Magnusson, M.S., 2002. T-patterns in behavior and DNA: detection and analysis with Theme and GeneTheme. *Proc. Measuring Behavior*, 160-161.
- Magnusson, M.S., 2004. *Theme* (Version 5.0) [Computer software]. Wageningen, NL: Noldus Information Technology.
- Magnusson, M.S., 2005. Understanding social interaction: discovering hidden structure with model and algorithms. In: Anolli, L., Duncan S. Jr., Magnusson, M.S., Riva, G. (Eds.), *The hidden structure of interaction: from neurons to culture patterns* (pp. 3-22). Amsterdam: IOS Press.
- Magnusson, M.S., 2006. Structure and communication in interactions. In: Riva, G., Anguera, M.T., Wiederhold, B.K., Mantovani, F. (Eds.), *From communication to presence: cognition, emotions and culture towards the ultimate communicative experience. Festschrift in honor of Luigi Anolli* (pp. 127-146). Amsterdam: IOS Press.
- Markov, A.A., 1971. Extension of the limit theorems of probability theory to a sum of variables connected in a chain. In: Howard, R. (Ed.), *Dynamic probabilistic systems, volume I: Markov models* (pp. 552-577). New York City: John Wiley & Sons. (Original work published 1906)
- Noton, D., Stark, L., 1971. Scanpaths in eye movements during pattern perception. *Science*, 171, 308-311.
- Pan, B., Hembrooke, H.A., Gay, G.K., Granka, L.A., Feusner, M.K., Newman, J.K., 2004. The determinants of web page viewing behavior: an eye-tracking study. *Proc. Eye Tracking Research and Applications (ETRA)*, pp. 147-154.
- Pieters, R., Rosbergen, E., Wedel, M., 1999. Visual attention to repeated print advertising: a test of scanpath theory. *Journal of Marketing Research*, 36, 424-438.
- Privitera, C.M., Stark, L.W., 2000. Algorithms for defining visual regions-of-interest: comparison with eye fixations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22, 970-982.
- Ross, S., 1996. *Stochastic processes* (2nd ed.). New York: John Wiley & Sons.
- Sankoff, D., Kruskal, J.B. (Eds.), 1983. *Time warps, string edits, and macromolecules: The theory and practice of sequence comparison*. Reading, MA: Addison-Wesley.
- Stark, L.W., Ellis, S.R., 1981. Scanpaths revisited: cognitive models direct active looking. In: Fisher, D.F., Monty, R.A., Senders, J.W. (Eds.), *Eye movements: Cognition and visual perception* (pp. 193-226). NJ: Lawrence Erlbaum.
- Stark Adam, P., Quinn, S., Edmonds, R., 2007. *Eyetracking the news: a study of print and online reading*. St. Petersburg, FL: The Poynter Institute for Media Studies.
- Smith, T.F., Waterman, M.S., 1981. Identification of common molecular subsequences. *Journal of Molecular Biology*, 147, 195-197.
- Unz, D.C., Werner, B., Mangold, R., Burmester, M., 2005. Information seeking in the WWW: Detecting T-patterns in eye movements and navigational behaviour. *Proc. Measuring Behavior*, Wageningen, The Netherlands.
- West, J.M., Haake, A.R., Rozanski, E.P., Karn, K.S., 2006. eyePatterns: software for identifying patterns and similarities across fixation sequences. *Proc. Eye Tracking Research and Applications (ETRA)*, pp. 149-154.

# A USABILITY STUDY OF MICROSOFT OFFICE 2007 AND MICROSOFT OFFICE 2003

Edgard Velez and Nestor J. Rodriguez  
*University of Puerto Rico - Mayaguez  
Mayaguez, Puerto Rico*

## ABSTRACT

Software vendors need to update their products periodically to improve functionality and add value to them. The expectations from the users' point of view are that the upgrades result in an improved interaction experience and better functionality and performance. These were the expectations of millions of users of the Office suit with the major changes in the user interface made to the 2007 version as compared to the 2003. The main objective of this study is to assess the effect of these changes in the usability of the Office applications. The study compares the usability of the 2003 and 2007 versions of the Word, Excel and Power Point applications based on five dependent variables: *completion time, number of tasks completed, errors committed, difficulty in performing the tasks, and user satisfaction*. The study followed a within subjects design. Fifty participants were asked to perform typical tasks in both versions of the three applications. All had at least one year of experience using the 2003 and 2007 versions of the three applications. The results of the study indicate that Excel 2003 is significantly superior to Excel 2007 in all the dependent variables. This supports the conclusion that the user interface of Excel 2007 worsened in comparison with the user interface of Excel 2003. However, it could not be concluded that the user interface of Word 2007 and Power Point 2007 improved in comparison to Word 2003 and Power Point 2003 respectively. Strictly from the point of view of usability, the users were left behind on the Office 2007 upgrade. They would have been better off keeping their 2003 version of the Office suit.

## KEYWORDS

Usability, user interfaces, office software, upgrades

## 1. INTRODUCTION

The advances in software and technology have changed the way we live and do everyday tasks. They enable us to perform tasks more quickly and effectively than ever before, and allow us to do things that seemed impossible in the past. One of the most used and widely distributed software in the world is MS Office. Word, Excel, and Power Point have become standards in their respective domain and are used by millions of users every day. Likewise, changes and updates to these systems will impact, positively or negatively, millions of users and can affect their effectiveness using the software, and their satisfaction.

A study by Michael (2003) concluded that updates to a program affect the learning process and require an adaptation and adjusting period for the users. Another related issue described in a study by Khoo (2006) was that upgrades could have a negative effect on user performance and also on the user satisfaction with the upgraded version. Similar studies conclude that the effect on the performance from upgrades or changes depends on the degree of changes made to the application (Lindgaard, 1991) and on the degree of change made on the features the users access the most (Nicholas 1990). Software upgrades, as described in the study by Shaw (2002), may also have a close relationship with user satisfaction.

The primary objective of this study is to determine the impact on user performance and satisfaction of the Office 2007 version compared to the 2003 version. In the upgrade to Office 2007 the icons organization, menu classification and navigation were dramatically modified, compared to the classic model displayed on all the previous versions of the applications. Such dramatic changes could potentially reduce the effectiveness, performance, and satisfaction of users. To have a better idea of how critical those changes made to the user interface could have affected the users performance, a user test is required that compares the aspects related to interface changes made in Office in 2007 with that of Office 2003. Thus, a usability test was performed to examine differences for time to complete a set of tasks, task completion rate, number of

errors committed, difficulty in performing the tasks and user satisfaction for users interacting with Word, Excel and PowerPoint. Statistical analyses were performed to determine significant differences.

A similar scope was followed in the study conducted by Beel (2007). For this study the participants were required to perform a series of tasks on Word 2003 and Word 2007. Before they performed the tasks they filled out a questionnaire about their experience with Word and demographic information. The completion time of each task on both versions was recorded. After the participants completed the tasks a second questionnaire was given to them to identify what kind of problems they ran into while performing the tasks, get their opinion on the transition from one version to the other, and determine which version they liked the most. The study concluded that the participants took more time to complete the tasks on the 2007 version. From the questionnaires, they concluded that most participants considered that they took more time to complete the tasks on the 2007 version because of the confusion created with the numerous changes on the interface. They expressed that in time they could get used to the new interface and they could be as efficient as they were with the previous version of the program

The study presented in this article, which is based on E. Velez's master thesis work (Velez, 2010), is broader than Beel's study. It involves two additional office applications and five dependent variables. In addition the participants had at least one year of experience using the applications.

## 2. EXPERIMENTAL DESIGN

The 2003 and 2007 versions of the Microsoft Office suite were used for the usability study. Following a within subjects methodology fifty participants were asked to perform several tasks on the Word, Excel and Power Point applications of each MS Office version. The sample consisted of students, professors and staff from a university. Each participant had at least one year of experience using Microsoft Office 2003 and 2007 versions and a basic knowledge of the Microsoft Windows operating system. They were selected on a first-come first-serve basis.

Before performing the tasks the participants were briefed about the test procedures and asked to sign consent forms. They were also asked to fill out a questionnaire to collect demographics data and information about their experience using computers and both versions of MS Office suits.

The tasks selected for the study were intended to make the participants interact with the features and options most commonly used for each application. These are listed in table 1. In order to compensate for the learning effect during the test, half of the participants conducted the tasks first on Office 2003 applications and the other half in Office 2007. To facilitate the data collection, the activity of the user interface was recorded using Morae, a tool for usability testing. This recording allowed to determine task completion times and analyze interaction sequences that resulted in unexpected or erroneous actions.

After performing the tasks the participants were asked to fill out a questionnaire to assess the level of difficulty they encountered performing each of the tasks and their satisfaction with each of version of each of the Office applications. These were measure using a 1-5 likert scale.

Table 1. Tasks performed on each application

Word	Excel	Power Point
Open file on desktop	Open file on desktop	Open file on desktop
Change font	Centralize title	Change slide design
Align paragraph	Insert row	Access slide master window
Draw a table	Delete a row	Change text color and add effects
Remove column from table	Calculate an average	Change bullets format
Add rows to the table	Replicate formula	Access slide sorter window
Insert image into document	Adjust decimal points	Add a new slide
Undo last action	Centralize data on cell	Draw and edit a circle
Add numbered bullets	Generate a graph	Insert an image into document
Run a Spell-check	0. Move graph	0. Insert sound
Run a search in the document	1. Assign title to cell	1. Run a slide show
Replace a word in the document	2. Copy data and paste	2. Print presentation handouts
Run a print preview	3. Highlight cell	
Save the document with a new name		
Copy content of one section		
Create a new document and paste the copied content		



The dependent variables of the study were *time to complete the task*, *number of tasks completed*, *number of errors committed*, *difficulty performing the tasks*, and *subjective user satisfaction*. The Dependent Samples *t* Test, was used to analyze *completion time* and the *number of error committed* while the Wilcoxon SUM Rank Test was used to analyze *user satisfaction*, *difficulty in performing the tasks*, and the *number of tasks completed*.

### 3. RESULTS

#### 3.1 Completion Times

A dependent-samples *t*-test did not reveal a significant difference between the average overall time it took the participants to complete all the tasks on Word 2003 (*Mean*=192.18, *SD*=73.84) and Word 2007 (*Mean*=198.54, *SD*=79.28). However, dependent sample *t*-tests revealed that there were significant differences in the average completion time for tasks 1, 4, 6, 10, 12, 13, 14 and 16. The participants were able to complete tasks 1 (*Open Document*), 10 (*Spell-Check*), 12 (*Text Replace*), 13 (*Preview Document*), 14 (*Save Document*) and 16 (*Create New Document*) significantly faster on Word 2003 than on Word 2007. On the other hand, the participants were able to complete tasks 4 (*Draw Table*) and 6 (*Add Rows*) significantly faster on Word 2007 than on Word 2003. Table 2 summarizes the average completion times for individual tasks of the three Office applications.

A dependent-samples *t*-test revealed a significant difference between the average overall time it took the participants to complete all the tasks on Excel 2003 (*Mean*= 170.04, *SD*= 103.09) and Excel 2007 (*Mean*= 217.50, *SD*= 102.55). The participants took less overall time to complete the tasks in Excel 2003 than in Excel 2007. Dependent samples *t*-test revealed that there were significant differences in average completion time for tasks 3, 6, 7, 8 and 9. The participants were able to complete tasks 3 (*Insert Row*), 7 (*Adjust Precision*), 8 (*Center Data*) and 9 (*Generate Graph*) faster on Excel 2003 than in Excel 2007. On the other hand, the participants completed Task 6 (*Replicate Formula*) faster on Excel 2007 than on Excel 2003.

A dependent-samples *t*-test did not reveal a significant difference between the average overall time it took the participants to complete all the tasks on PowerPoint 2003 (*Mean*= 286.26, *SD*= 111.29) and PowerPoint 2007 (*Mean*= 285.38, *SD*= 95.83). However, there were significant differences in some of the individual tasks. Dependent sample *t*-test revealed that there were significant differences in the average completion time for tasks 1, 2, 3, 5, 6, 7, 8 and 10. The participants were able to complete tasks 1 (*Open Document*), 2 (*Change Design*), 3 (*Slide Master*), 5 (*Change Bullets*), 6 (*Slide Sorter*) and 7 (*Add Slide*) faster on PowerPoint 2003 than on PowerPoint 2007. On the other hand, the participants were able to complete task 8 (*Draw / Edit Circle*) and 10 (*Insert Sound*) faster on Power Point 2007 than in 2003.

Table 2. Average completion times for individual tasks in 2003 and 2007 versions of Office applications

Task	Word 2003		Word 2007		Excel 2003		Excel 2007		P. Point 2003		P. Point 2007	
	Mean (sec)	SD (sec)	Mean (sec)	SD (sec)	Mean (sec)	SD (sec)	Mean (sec)	SD (sec)	Mean (sec)	SD (sec)	Mean (sec)	SD (sec)
1	4.8	1.4	5.7	1.4	5.1	5.8	5.4	2.1	5	1.6	6.3	2.2
2	13.3	5.9	13.5	4.4	37.2	43.6	33.8	39.9	19.9	17.9	30.4	22.7
3	3.7	1.1	3.9	1.4	4.7	3.8	8.9	14.9	50.9	38.9	64.6	35.1
4	22.9	14.0	13.4	5.4	5.1	6.3	4.6	4.7	61.8	38.5	55.6	27.9
5	5.5	5.6	7.0	8.0	26.0	27.0	24.8	18.0	12.2	7.9	16.7	14.4
6	22.3	28.0	11.8	10.6	15.5	30.4	9.4	13.8	9.6	7.1	14.6	14.8
7	9.7	5.4	10.0	9.0	11.6	13.3	18.4	28.3	5.4	4.6	7.78	7.9
8	1.2	0.5	1.2	0.5	7.4	2.9	8.3	4.5	46.8	33.8	31.8	15.3
9	31.5	12.0	32.5	14.7	33.7	13.4	77.6	33.6	13.5	6.1	16.0	12.8
10	22.1	15.3	28.4	21.0	4.0	3.1	3.6	2.4	48.3	32.9	28.9	23.8
11	14.1	17.1	16.4	20.4	4.9	2.4	5.1	1.8	3.0	5.6	3.2	4.3
12	16.8	12.8	20.8	19.7	8.4	4.5	10.0	6.1	8.9	8.9	9.5	4.8
13	2.9	2.2	7.7	7.5	6.6	4.7	7.6	5.5				
14	11.6	3.1	13.3	4.5								
15	5.8	2.6	6.23	2.6								
16	3.8	3.4	6.7	3.8								

### 3.2 Number of Tasks Completed

Table 2 summarizes the number of participants that completed each task performed on each version of the Office applications. Wilcoxon signed ranks tests did not reveal significant differences in the total number of tasks completed by the participants on Word. The participants completed an average of 15.78 tasks on Word 2003 and also 15.78 tasks on Word 2007. No significant differences were found in task completion for individual task on Word 2003 and Word 2007.

Table 2. Number of participants that complete each task performed on each version of the Office applications

Task	Word 2003	Word 2007	Excel 2003	Excel 2007	P. Point 2003	P. Point 2007
1	50	50	50	50	50	50
2	50	50	40	40	47	45
3	50	50	50	50	38	31
4	50	50	50	50	36	38
5	50	50	47	47	50	49
6	47	49	46	44	49	48
7	50	49	48	47	50	50
8	50	50	50	50	44	47
9	48	49	48	27	49	48
10	50	48	50	50	37	49
11	48	47	50	50	50	50
12	49	48	50	50	49	49
13	50	49	50	50		
14	50	50				
15	50	50				
16	50	50				

A Wilcoxon signed ranks test revealed significant difference in the total number of tasks completed by the participants on Excel. The participants completed an average of 12.58 tasks on Excel 2003 and 12.10 on Word 2007. For individual tasks the Wilcoxon signed ranks test revealed a significant difference only for Task #9 (*Generate Graph*). Almost all the participants (48) were able to complete Task #9 in Excel 2003 while almost half (27) of the participants were able to complete it in Excel 2007.

No significant difference in the total number of tasks completed was found between the two versions of PowerPoint. The participants completed an average of 10.98 tasks on PowerPoint 2003 and 11.08 tasks on PowerPoint 2007. The only task that showed a significant difference was Task #10 (*Insert Sound*). Almost all the participants completed this task on Power Point 2007 while only 37 completed it on Power Point 2003.

### 3.3 Number of Errors Committed

In the context of the test an erroneous action in performing a task was a deviation from an expected sequence(s) of steps to properly execute it. Dependent sample *t* tests revealed that there were significant differences in the overall average number of errors committed by the participants on the three applications of Office. The participants committed significantly fewer errors on Word 2007 (*Mean*= 1.12) than on Word 2003 (*Mean*= 2.02). However, the participants committed significantly fewer errors on Excel 2003 (*Mean*= 2.20) than on Excel 2007 (*mean*= 3.28). In addition the participants committed significantly fewer errors on Power Point 2003 (*Mean*= 2.26) than on Power Point 2007 (*mean*= 3.06).

### 3.4 Difficulty in performing the Tasks

The difficulty in performing the tasks was measured with a post-test questionnaire that asked the participants to rate the difficulty in performing each task using a scale from 1 to 5, with 1 meaning very difficult and 5 very easy. Table 4 summarizes the average scores for each task on each version of the Microsoft Office applications. Wilcoxon tests did not reveal a significant difference in the average overall difficulty in performing the tasks between the two versions of Word. The overall average difficulty in performing the

tasks rating for Word 2003 was 4.80 while for Word 2007 was 4.82. For individual tasks Wilcoxon tests revealed significant differences in user difficulty in performing tasks 4, 6 and 16. The participants had less difficulty on Word 2007 than on Word 2003 on Task 4 (*Draw a table*) and Task 6 (*Add rows to the table*), while the opposite resulted with Task 16 (*New Document*).

Table 4. Average scores for difficulty in performing each task on each version of the Office applications

Task	Word 2003	Word 2007	Excel 2003	Excel 2007	P. Point 2003	P. Point 2007
1	5.00	4.98	5.00	4.94	5.00	4.86
2	4.94	4.96	3.88	3.84	4.26	4.00
3	4.90	5.00	4.80	4.80	3.48	3.08
4	4.44	4.80	4.86	4.92	4.58	4.46
5	4.72	4.82	4.42	4.38	4.82	4.72
6	4.34	4.80	4.62	4.62	4.66	4.50
7	4.78	4.76	4.56	4.38	4.94	4.86
8	5.00	4.96	4.94	4.84	4.24	4.34
9	4.70	4.70	4.60	3.84	4.76	4.80
10	4.80	4.66	4.92	4.92	3.44	3.94
11	4.64	4.60	5.00	4.80	4.94	4.88
12	4.60	4.58	4.80	4.72	4.84	4.66
13	4.90	4.68	4.90	4.92		
14	4.98	4.92				
15	4.98	5.00				
16	5.00	4.90				

A Wilcoxon tests revealed significant difference in the average overall difficulty in performing the tasks between the two versions of Excel. The participants had less difficulties with Excel 2003 than with Excel 2007. The overall average difficulty score was 4.71 for Excel 2003 and 4.60 for Excel 2007. Wilcoxon tests also revealed significant differences in user difficulty in performing Tasks 9 (*Generate Graph*) and Task 11 (*Title Cell*). For both tasks the participants had less difficulties with Excel 2003 than with Excel 2007.

Wilcoxon tests did not reveal a significant difference in the average overall difficulty in performing the tasks between the two versions of PowerPoint. The overall average difficulty score was 4.51 for PowerPoint 2003 and 4.43 for PowerPoint 2007. Significant differences in user difficulty performing the tasks were found for Task 1 (*Open Document*) and Task 12 (*Print Handouts*). For both tasks the participants had less difficulties on PowerPoint 2003 than on PowerPoint 2007.

### 3.5 Satisfaction

The participants were asked to rate their satisfaction with each Office application using a scale from 1 to 5, with 1 meaning very dissatisfied and 5 very satisfied. Wilcoxon tests did not reveal significant differences in the participants' satisfaction with Word and Power Point applications. The overall satisfaction rating for Word 2003 was 3.98 while for Word 2007 was 4.20. The overall satisfaction rating for PowerPoint 2003 was 3.54 while for PowerPoint 2007 was 3.84. A Wilcoxon test did reveal a significant difference in the overall satisfaction rating on Excel. The overall satisfaction rating for Excel 2003 was 4.24 while for Word 2007 was 3.70.

## 4. DISCUSSION OF RESULTS

An analysis of the video recordings was performed for individual tasks that resulted in significant differences for the dependent variables on the two versions of the applications. This analysis revealed that the significant differences were mainly due to *retroactive interference* (new information causing the loss of old information), *proactive inhibition* (old information interfering with new information), shuffling of options from menus and tabs, elimination or addition of short cuts, and clustering of icons. Following are some examples of them.

*Retroactive interference* – The procedure for inserting a table in Word 2007 is to select the **Insert** menu and choose the table icon, which is a simpler procedure than the one needed to execute the task in Word 2003. When performing this task in Word 2003 some users struggle recalling the procedure for this version.

*Proactive inhibition* – The users can change bullets in Power Point 2003 by selecting the **Format** menu and choosing the **Bullets** option. Although this option is no longer available in the **Format** menu of Power Point 2007, many users attempted to change bullets by selecting this menu.

*Shuffling options from menus and tabs* - In Word 2003 the spell-checking task can be accomplished by selecting the **Tools** menu and then choosing the **Spelling & Grammar** option. In Word 2007 the **Tools** menu was eliminated and the spell checking function was moved to the **Review** tab. This change took a toll on the participant's performance because many of them had a hard time remembering where the function was moved in Word 2007.

*Addition of short cuts* – In Word 2007 the users can add rows to a table by selecting that option with a mouse right-click short cut. Even though this short cut is not available in Word 2003, many users tried to add rows by attempting this option.

*Cluttering of icons* – The task to access the slide master required accessing the **View** menu and choosing the **Slide Master** option in Power Point 2003. Similarly, in Power Point 2007 the user selects the **View** tab and choose the **Slide Mater** icon. In performing this task in Power Point 2007 many of the participants accessed the right tab, but because many of the icons in that tab were similar, they had trouble identifying the correct one or went looking for it elsewhere.

## 5. CONCLUSION

The results of the study revealed significant differences between Excel 2003 and 2007 for all dependent variables. The results also revealed a significant difference between the two versions of Word and the two versions of Power Point for the dependent variable *errors committed*. Table 5 summarizes these results. Shaded squares indicate the superior version for the particular dependent variable.

Table 5. Significant differences for dependent variables between 2003 and 2007 Office applications

Application	Completion Times	Tasks Completed	Errors	Difficulty performing the tasks	Satisfaction
Word			2007		
Excel	2003	2003	2003	2003	2003
PowerPoint			2003		

The results of the study indicate that the participants did significantly better on Excel 2003 than Excel 2007 for all dependent variables. The results also indicate that there were significant differences in terms of errors committed by the participants in the three Office applications. The participants performed significantly better on Excel 2003 and PowerPoint 2003. However, they performed significantly better on Word 2007.

These results support the conclusion that the user interface of Excel 2007 worsened in comparison with the user interface of Excel 2003. However, although the participants committed significantly less error on Power Point 2003, it cannot be concluded that the user interface of this version is superior in terms of usability to that of Power Point 2007. Similarly, it cannot be claimed that Word 2007 is superior to Word 2003.

Taking in to account the technological advances, the advances made in the design of user interfaces, and the investment made by the developer in the upgrade, the logical expectation would has been for table 5 to be predominantly shaded with superiority for the 2007 version of the Office suit. A more conservative outcome would have been no change in usability for the upgrade. Thus, strictly from the point of view of usability, the users would have been better off keeping their 2003 version of the Office suit. Investing in an upgrade from Office 2003 to 2007, and now 2010, is something to consider if the new capabilities of the upgrade are worth the investment. The vendor should take note on these results because they reflect that the users were left behind on the Office 2007 upgrade.

## REFERENCES

- Beel C., 2007. Analysis of Microsoft Office 2007 User Interface Design. *Proceedings of the 5th Winona Computer Science Undergraduate Research Seminar*. Winona, MN.
- Khoo, H.M., 2006. *Upgrading Packaged Software: An Exploratory Study of Decisions, Impacts, and Coping Strategies from the Perspectives of Stakeholders*.(2006). CIS Dissertations. Paper 5.  
[http://digitalarchive.gsu.edu/cis\\_diss/5](http://digitalarchive.gsu.edu/cis_diss/5)
- Lindgaard, G., Bednall, E. and Chessari, J., 1991. Improving User performance. *IEEE Journal on Communications*. Vol. 9, Issue 4, pg. 506-517.
- Michael, C.L. et al, 2003. Adoption Process of Upgrading Software: an Empirical Study of Windows XP. *Proceedings of the 36th Annual Hawaii International Conference on Systems Sciences*. Clear Water Bay, China, pp. 9.
- Nicholas, P. and O'Connor, M., 1990. Variations in User Satisfaction with MIS. *Proceedings of the Twenty-Third Annual Hawaii International Conference on System Sciences*. Kailua-Kona, HI, pp. 479 - 487.
- Shaw, N.C., DeLone, W.H. and Niederman, F., 2002. Sources of Dissatisfaction in End-user Support: An Empirical Study. *ACM SIGMIS Database*, Vol. 33, Iss. 2; pg. 41.
- Velez, E., 2010. A Comparative Usability Study of Microsoft Office 2007 and Microsoft Office 2003, Master Thesis, University of Puerto Rico – Mayaguez.

# EVALUATION OF DIFFERENT INTERACTION TECHNIQUES FOR TOUCH DEVICES

Michael Raschke<sup>1</sup>, Tanja Blascheck<sup>1</sup>, David Bold<sup>1</sup>, Dominik Muhler<sup>1</sup>,  
Thomas Ertl<sup>1</sup> and Thomas Schlegel<sup>2</sup>

<sup>1</sup>*Institute for Visualization and Interactive Systems, University of Stuttgart  
Universitätsstraße 38, 70569 Stuttgart, Germany*

<sup>2</sup>*Institute of Software- and Multimedia-Technology, Technische Universität Dresden  
Nöthnitzerstraße 46, 01062 Dresden, Germany*

## ABSTRACT

New interaction techniques beyond mouse and keyboard were established during the last five years allowing different and multiple modalities. The aim of our work was to study how users work with a touch device, which can be used as a basis for designing natural user interfaces optimized for touch. We developed an evaluation concept for a single- versus multi-touch scenario. Applying this concept we studied users' strategies, accuracy and completion times in performing the three evaluation tasks: moving virtual objects by touching, typing with a virtual keyboard and using touch menus to select options. The user tests were performed on a Microsoft Surface during the open day at our university. Results of the evaluation showed that our two user groups "beginners" and "professionals" directly profit from the natural interaction concepts of a touch interface, like for example direct manipulation, except for scenarios where a haptic feedback is helpful to minimize input errors. Both groups only differ in using a virtual keyboard due to training effects. Additionally, the intuitive usage and novelty of a touch device seemed to motivate the participants in accomplishing the tasks.

## KEYWORDS

Human-Computer-Interaction, Touch, Microsoft Surface, Evaluation, Natural User Interface

## 1. INTRODUCTION AND RELATED WORK

Today, the possibilities of interacting with computer systems completely differ from what they were like at the beginning of the computer era. Most users interact with personal computers using a mouse and a keyboard. Together with graphical user interfaces (GUIs) these two interaction devices have been the state of the art computer interfaces during the last 20 years. This is changing now. New interaction techniques beyond mouse and keyboard like touch-based systems profit from new technologies and are now massively used in consumer products. Touch based systems are often described with the term natural user interfaces (NUI 2011), also because of their direct manipulation interaction concept with easy and stringent hand-eye coordination. In our work we therefore concentrated on interaction techniques for touch based systems to strengthen and facilitate interaction on the basis of touch-based NUIs.

The aim of our evaluation was to study how users work with a multi-touch device, which can be used as a basis for designing NUIs optimized for touch. We wanted to know how beginners' behavior differs from that of professionals in their strategies, accuracy rates and completion times. Another aim was the development of an evaluation process for studying the usability of touch devices in a generalized way. One major point is the definition and differentiation of a single- versus a multi-touch scenario as a first dimension of describing touch interaction and the usage of touch-enabled devices by a single person or by several persons (single-user versus multi-user scenario) as a second dimension.

In this work we describe main evaluation criteria which can be used to perform a touch-study, followed by our study procedure, the study results and a discussion about future aspects and possible developments.

For our study tests we used the Microsoft Surface (see figure 1). The Microsoft Surface is a software and hardware combined multi-touch sensitive device which allows users to interact with an integrated PC by using gestures. The screen has a 76 cm reflective surface with a DLP projector in XGA resolution (1024x768

pixels). It is positioned underneath the screen and projects an image on its transmissive surface from below. Reflections of infrared light from objects and human fingertips on the surface are recognized by five cameras inside the device. The Surface can respond to up to 52 touches at the same time. The device runs Windows Vista by default. More information about the Microsoft Surface is available under (Surface 2011).



Figure 1. The Microsoft Surface was used for the evaluation of different interaction techniques for touch devices.

Müller-Tomfelde et al. presented evaluation results on commonalities and differences between touch and mouse input for co-located interaction between teams of two people (Müller-Tomfelde et al. 2008). Kammer et al. described a test environment within they showed the applicability of their formalization approach of gestural interaction on multi-touch surfaces (Kammer et al. 2010). Gross et al. conducted a user evaluation to study a pure multi-touch paradigm where no single-touch input was possible (Gross et al. 2008). A study about the effects of co-present embodiments on awareness and collaboration in tabletop groupware was conducted by Pinelle et al. 2008. They determined whether factors such as size, realism, and visibility can improve awareness and coordination. The differences between direct-touch and mouse input for tasks on tabletop displays were investigated by Forlines et al. 2007. With a user study Lucchi et al. compared touch and tangible interfaces (Lucchi et al 2010).

## 2. EVALUATION CRITERIA

We identified twelve interaction concepts for touch devices and divided them into three main groups (see table 1). These three main groups describe to which interaction elements they are applied, either to graphical primitives, to input techniques or to view port handling. Graphical primitives are for example rectangles, circles etc. but also more complex grouped vector graphics. These graphical elements can be created, selected, translated, scaled or rotated. The application to input techniques covers the input of text, numerical values but also the selection of an option. View port handling is possible in two and three dimensions. Each interaction technique can be realized by direct manipulation of the touch screen element with finger touches, with a general gesture or with an input of values in input fields of a dialogue.

Table 1. Classification of interaction concepts into three groups and their application to different interaction elements.

<b>Application on:</b>	<b>graphical primitives</b>	<b>input techniques</b>	<b>view port handling</b>
<b>Interaction concept:</b>	<ul style="list-style-type: none"> <li>• Creation</li> <li>• Manipulation and single Selection</li> <li>• Hierarchical selection</li> <li>• Multi selection</li> <li>• Translation</li> <li>• Scaling</li> <li>• Rotation</li> </ul>	<ul style="list-style-type: none"> <li>• Input of text</li> <li>• Input of numerical values</li> <li>• Selecting an option</li> </ul>	<ul style="list-style-type: none"> <li>• Viewport shifting in two dimensions</li> <li>• Viewport shifting in three dimensions</li> </ul>

We defined a relationship between interaction concepts, the numbers of users using this concept and the interaction mode with the device, either single-touched or multi-touched. A multi-user scenario exists, if several users are working with the same interface. “Single-touch” means that the user is only interacting with the interface by using one finger. “Multi-touch” means the utilizing of more than one finger, e.g. several fingers of one hand or both hands. The relationship between interaction concepts, user numbers and interaction modes are visualized in the matrix shown in figure 2.

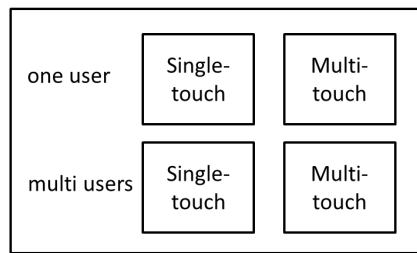


Figure 2. Relationship between interaction concepts, user numbers and interaction modes.

We defined two different user groups following the approach of (Thaller 2002): “beginners” and “professionals”. Beginners are people who are working with computers for less than 10 hours per week and have nearly no touch device experience. Professionals are using touch devices for more than 10 hours per week and use touch devices frequently.

Based on the classification of interaction concepts and the relationship matrix of figure 2 we identified five general interaction techniques with which a task can be accomplished in a multi-touch environment: direct manipulation, option selection, gestures, dialogues and forms, and an iconic representation of possible interactions and software actions.

To study these interaction concepts we defined the following user study tasks:

*Task 1:* Move virtual objects on the screen.

*Task 2:* Type a given text with a virtual keyboard.

*Task 3:* Select a given option in a menu.

Task 1 includes the interaction techniques direct manipulation and gestures. Task 2 covers filling out dialogues or forms. Task 3 studies the interaction technique option selection. For this task we compared the classical menu elements and the Microsoft Surface element menu (see figure 3). For all tasks we measured accuracy rates and completion times.

### 3. USER STUDY

The aim of our evaluation was to proof the viability of the evaluation concept and criteria described in section two. Additionally all tasks studied differences between the two user groups. The user study was conducted with 34 participants during the open house at our university in the summer of 2010, which gave us the possibility to approach a variety of different users. This section describes the subject properties, the experimental setup and the study procedure. Results will be presented and discussed in section four.

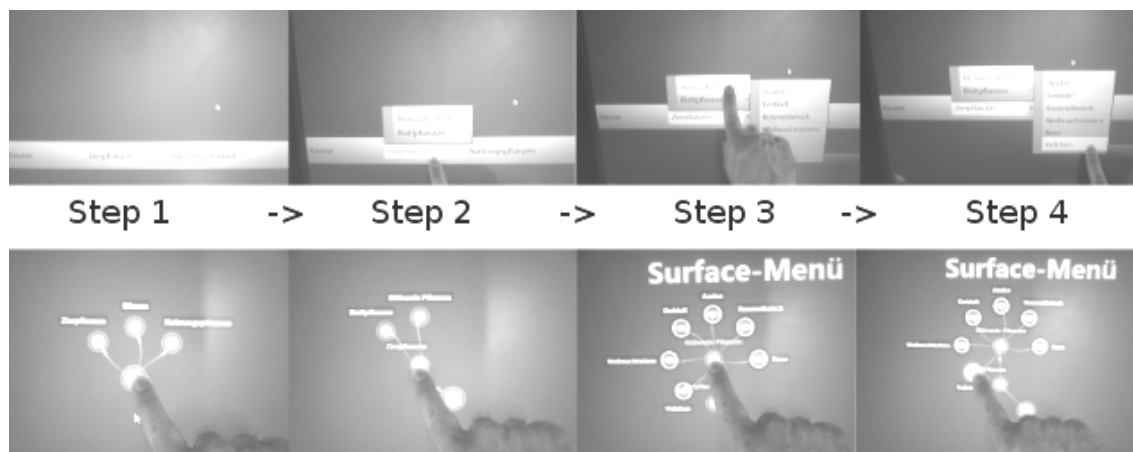


Figure 3. Interaction technique “option selection” in four steps realized by classical menu elements and by the Microsoft Surface element menu.



**Subjects:** Thirty-four participants (average age 21 years, minimum 8 years, and maximum 67 years). Sex was not considered a confounding factor for this study. Subjects were not paid for their participation in the user experiment. We didn't perform a color vision test or a visual test.

**Experimental Setup:** We developed three applications for the evaluation with the Microsoft Surface SDK (see figure 4 a-d). The user experiment was conducted in a lecture hall in which other scientific groups also had stands and presented their work. Thus, the subjects participated in a very public fashion. We placed the Microsoft Surface centrally at our stand to have a lot of space around it. To make the experiment as comfortable as possible we placed some chairs around the Surface. This public evaluation scenario corresponds to a typical operational domain for this device.

**Study Procedure.** At the beginning, subjects had to fill out a multiple-choice questionnaire about their age, field of study or profession, computer skills and prior knowledge of touch devices. They also had the opportunity to state their expectations regarding multi-touch devices in an open question. Then, they were generally introduced to the Microsoft Surface and the study procedure. The main evaluation took between 10 and 15 minutes depending on the knowledge of the subjects. Every part of the user study was started with a short introduction about the specific task. There was a "Give Up" option available – but this was not used by the subjects.

*Task 1:* First, subjects had to move 22 small rectangles on the Microsoft Surface screen (figure 4a). Each half of them were colored in red and blue. The aim of the task was to move them correctly inside a big blue or red filled out area on the screen. The rectangles were randomly positioned and initially none of them was positioned inside the boxes. Subjects were free to use one finger, two fingers, only one hand or both hands. Task completion times and accuracy rates were measured. The movement of an object was counted as an error, if the user released an object or moved it to the wrong corner. Completion times were measured.

*Task 2:* Second, we evaluated the virtual keyboard (Figure 4b). Participants had to type the sentence "Rosen sind rot, Veilchen sind blau, doch daraus werde ich nicht schlau." (German for „Roses are red, Violets are blue, sugar is sweet and so are you.“) in two ways. The first way was to type the poem as fast as possible without paying attention to typing errors. The second way was to type the sentence with a minimum of typing errors but without paying attention to the completion time. At the end we counted the typing errors. Additional, every unnecessary typing was counted as an error. We measured the completion times for both ways.

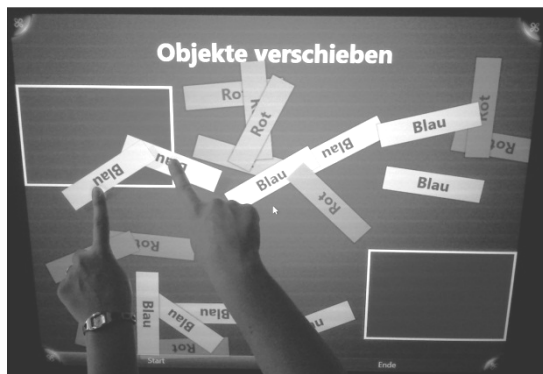


Figure 4a). The aim of task 1 was to correctly move blue and red colored rectangles inside a blue (top left corner) or red (bottom right corner) filled area. The colored rectangles were marked with the tags "Rot" and "Blau" (german for red and blue).

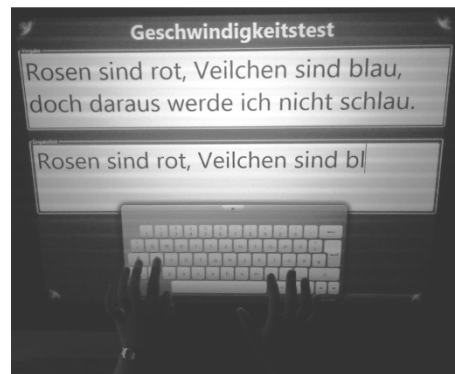


Figure 4b). Evaluation of the virtual keyboard in task 2. Participants had to type the sentence "Rosen sind rot, Veilchen sind blau, doch daraus werde ich nicht schlau." (German for „Roses are red, Violets are blue, sugar is sweet and so are you.“).

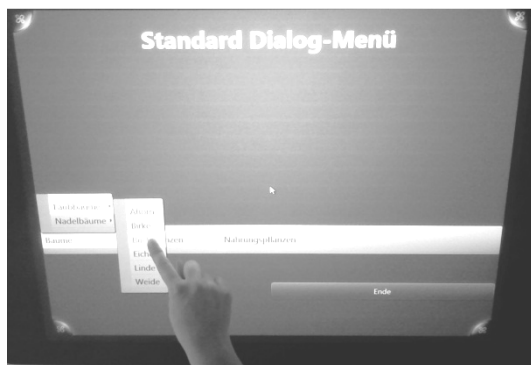


Figure 4c). Task 3 evaluated classical menu elements

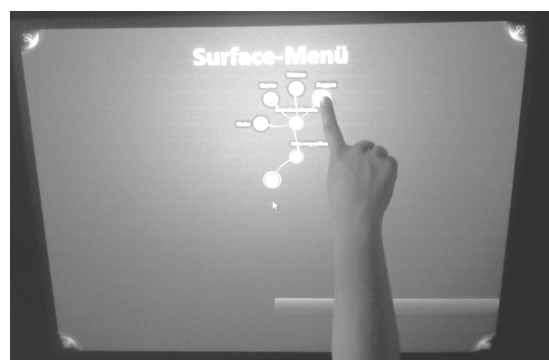


Figure 4d). versus the Surface menu element.

*Task 3:* The third task evaluated classical menu elements (figure 4c) versus the Microsoft Surface menu element (figure 4d). In both cases subjects had to find a specific element in the menu that was presented by the operator. The menu structure was the same for both menu types (see also figure 3). This task was conducted counter balanced between subjects. Completion times and accuracy rates were measured. We identified errors when subjects released the menu before reaching the final menu element or not following the right path through the menu.

After the main evaluation, participants were given a second questionnaire, in which they ranked in a Likert scale the usability of the onscreen keyboard, the two menu types, how they perceived the touch interface in general and the self-descriptiveness of the GUI-elements. Finally, we asked them whether their expectations of a multi-touch device were fulfilled or not.

#### 4. STUDY RESULTS AND DISCUSSION

The results of our first questionnaire showed that thirteen subjects are using computers for less than ten hours per week, 21 for more than ten hours per week. Fourteen participants stated that they don't use devices with touch interfaces. Twenty participants stated that they either use touch mobiles, smartphones, navigation devices or copiers (multiple answers were possible). Thirteen subjects use touch devices for less than one hour per week, four between one and two hours and three for more than two hours. Five participants also use multi-touch devices. Three of them use these devices for less than two hours, two between four and ten hours per week. Most of the subjects expected from a multi-touch device fun and an easy, intuitive as well as self-explanatory interaction. Two of the thirteen subjects who are using computers for less than ten hours per week use touch devices (mostly navigation devices). This percentage is the same for the user group who uses computers for more than ten hours per week.

##### Study Results

Table 1 shows that completion times only differ in typing with the keyboard. For both user groups the median of the task 3 completion times "select a given option in a menu" show that users accomplish task 3 a little bit faster with classical menu elements than with the surface menu element.

Table 1. Median of completion times for the tasks: correctly move objects, type a sentence (with and without paying attention to accuracy) and select an option with the classical menu elements and with the Microsoft Surface menu element. This table presents the results with a differentiation between the use times of touch systems (SD = standard deviation).

	T1: Moving virtual objects	T2: Keyboard (without paying attention to accuracy)	T2: Keyboard (with paying attention to accuracy)	T3: Classical Menu	T3: Microsoft Surface Menu
Less than 10 hours (in seconds)	22 (SD 6,17)	45 (SD 33,3)	60 (SD 14,8)	5 (SD 4,1)	7,5 (SD 13,3)
More than 10 hours (in seconds)	20 (SD 5,5)	28,5 (SD 18,6)	46 (SD 17,1)	5 (SD 3,8)	7 (SD 6,0)

Table 2 shows that the accuracy rates for all tasks are approximately equal. The accuracy rates only differ when moving virtual objects or typing with the keyboard, if paying attention to accuracy.

Table 2. Median of accuracy rates (counted errors) for the tasks: correctly move objects, type a sentence (with and without paying attention to accuracy) and select an option with the classical menu elements and with the Microsoft Surface menu element. This table presents the results with a differentiation between the use times of touch systems (SD = standard deviation).

	T1: Moving virtual objects	T2: Keyboard (without paying attention to accuracy)	T2: Keyboard (with paying attention to accuracy)	T3: Classical Menu	T3: Microsoft Surface Menu
Less than 10 hours	2 (SD 1,6)	11 (SD 12,7)	4 (SD 4,5)	0	0
More than 10 hours	4 (SD 2,4)	8,5 (SD 6,6)	10 (SD 7,4)	0	1 (SD 1,9)

There are no significant differences between users with touch experience and without touch experience, neither in the completion times nor in the accuracy rates. This issue is shown in table 3 and table 4.

Table 3. Median of completion times for the tasks: correctly move objects, type a sentence (with and without paying attention to accuracy) and select an option with the classical menu elements and with the Microsoft Surface menu element. This table presents the results with a differentiation between the user experiences with touch systems (SD = standard deviation).

	T1: Moving virtual objects	T2: Keyboard (without paying attention to accuracy)	T2: Keyboard (with paying attention to accuracy)	T3: Classical Menu	T3: Microsoft Surface Menu
Without experience (in seconds)	21,5 (SD 5,3)	35 (SD 31,8)	51,5 (SD 18,8)	6,5 (SD 4,5)	8,5 (SD 11,4)
With experience (in seconds)	20 (SD 6,3)	32 (SD 21,7)	48 (SD 15,0)	4 (SD 2,9)	6 (SD 7,2)

Table 4. Median of accuracy rates (counted errors) for the tasks: correctly move objects, type a sentence (with and without paying attention to accuracy) and select an option with the classical menu elements and with the Microsoft Surface menu element. This table presents the results with a differentiation between the user experiences with touch systems (SD = standard deviation).

	T1: Moving virtual objects	T2: Keyboard (without paying attention to accuracy)	T2: Keyboard (with paying attention to accuracy)	T3: Classical Menu	T3: Microsoft Surface Menu
Without experience	3,5 (SD 2,3)	10 (SD 9,1)	7,5 (SD 4,1)	0,5 (SD 0,8)	0
With experience	3 (SD 2,2)	11 (SD 9,5)	6 (SD 8,5)	0	1 (SD 1,2)

With the second questionnaire we asked for the usability of the onscreen keyboard (table 5, questions a and b), the two menu types (table 5, question c), users' touch feelings in general (table 5, questions d and e).

Table 5. Likert scaled results of the second questionnaire about usability of the onscreen keyboard, the two menu types and users' touch feelings in general (SD = standard deviation).

Question	Results
a) I found typing with the virtual keyboard motivating.	4,8 (1 frustrating, 6 motivating) (SD 1,2)
b) I found typing with the virtual keyboard usual.	2,9 (1 unusual, 6 usual) (SD 1,4)
c) Which menu style do you prefer?	1,8 (1 classical menu elements, 2 surface element menu) (SD 0,5)
d) I found using the multi-touch device motivating.	4,7 (1 frustrating, 6 motivating) (SD 1,3)
e) I found using the multi-touch device usual.	2,8 (1 unusual, 6 usual) (SD 1,6)

Twenty-four participants stated that their expectations of a multi-touch device were fulfilled.

#### Discussion of Results

Both user groups don't differ significantly in performing task 1 and task 3 because experience with touch devices is not basically necessary for accomplishing these tasks quickly. Task 1 is very similar to a typical real world movement scenario. Humans often move real objects on the surface of a table, for example papers or pencils. The eye-hand coordination for object movements in a real world scenario is different from moving a mouse-device over a short distance which indirectly moves a mouse-pointer on the screen. However, while the movements on the table surface are the same as the movements on a touch surface, a haptic feedback is missing. Users stated that they are missing a haptic feedback which might explain the

errors, e.g. unintentionally releasing an object. Generally it appears that users tend to trigger actions due to unintended touches in areas outside the focus area. Similar to (Gross et al. 2008) we found no differences between users who are familiar with PCs and used the Microsoft Surface for the first time.

Task 2 is a typical scenario where users profit from practice. The differences between beginners' and professionals' speed are well measured in typical keyboard typing tests. Our results suggest that the results of this task can directly be transferred to a virtual screen keyboard scenario. We observed during the tests that professionals type faster because they are accustomed to work with the QWERTY-keyboard. It is interesting, that only one participant touched directly into the text to correct it while most of the others used the backspace key or cursor keys. There are possible explanations for this behavior:

- Users are focused on the keyboard and ignore other possibilities.
- They don't use a pointing device to correct a text but rather use discrete navigation (e.g. backspace).
- They are not used to positioning and editing by direct touch manipulation in the text field, although some participants moved the virtual keyboard by touch.

Because there are only differences in completion times and accuracy rates for the keyboard task this evaluation shows that users directly profit from the natural interaction concepts of a touch interface, like for example direct manipulation except for scenarios where a haptic feedback is helpful to minimize input errors.

The intuitive usage and novelty of a touch device seems to motivate the participants in accomplishing the tasks (table 5, questions a and d) even in using the virtual keyboard. The novelty can directly be seen from the unusualness (table 5, questions b and e). The Surface element menu was probably preferred due to its novelty and to a behavioral deviation in the implementation of the classical menu.

## 5. CONCLUSION AND FUTURE WORK

In this paper we have presented evaluation criteria of different interaction techniques with touch devices for different user groups. We conducted a user study with which we studied the following interaction techniques: moving objects on the surface, typing with the virtual keyboard, using menus to select options.

The participants of our study were separated into the two user groups "beginners" and "professionals". We measured completion times and accuracy rates. All results showed that there are no significant differences between beginners and professionals with the exception of typing with the virtual keyboard. Our results suggest that results from classical keyboard tests can directly be transferred to a virtual screen keyboard scenario, except interesting issues of text correction. In conclusion, our evaluation showed that users directly profit from the natural interaction concepts of a touch interface except for scenarios where a haptic feedback is helpful to minimize input errors.

During the development of our evaluation we had comprehensive discussions about the transfer of well-known usability criteria from the legacy "mouse and keyboard world" to the "touch world". In an ongoing project we study the applicability of the ISO 9241-110 (ISO 2010) norm's criteria "suitability for the task", "suitability for learning", "suitability for individualization", "conformity with user expectations", "self-descriptiveness", "controllability", and "error tolerance" to touch GUIs. We extend the ISO 9241-110 norm with our evaluation criteria single- versus multi-user and single- versus multi-touch interaction. Our goal is to define a global concept that comprises all these aspects, based on the findings in this evaluation. The study procedure described in this paper is mainly focused on single-user scenarios. To get a better insight into the ongoing questions we are planning a scenario that compares single- and multi-user interactions on a multi-touch device and use a head mounted eye tracking system as well as an image recognition system. Based on these results we will develop cognitive based algorithms which will be applied to optimize the layout of a graphical touchable user interface.

## ACKNOWLEDGEMENT

We would like to thank the many participants for taking part in our user study.

## REFERENCES

- Forlines, Clifton et al, Direct-touch vs. mouse input for tabletop displays, 2007. *Proceedings of the SIGCHI conference on Human factors in computing systems*. San Jose, USA, pp. 647-656.
- Gross, Tom et al, 2008. The cuetable: cooperative and competitive multi-touch interaction on a tabletop. *CHI '08 extended abstracts on Human factors in computing systems*. Florence, Italy, pp. 3465-3470.
- ISO: International Organization for Standardization, ISO 9241-110, <http://www.iso.org>, last retrieved February, 2011.
- Kammer, Dietrich et al, Towards a Formalization of Multi-touch Gestures, 2010. *ACM Interactive Tabletops and Surfaces 2010*. Saarbrücken, Germany, pp. 49-58.
- Lucchi, Aurélien et al, An empirical evaluation of touch and tangible interfaces for tabletop displays, 2010. *Proceedings of the fourth international conference on Tangible, embedded, and embodied interaction*. Cambridge, USA, pp. 177-184.
- Müller-Tomfelde et al, Touchers and mousers: commonalities and differences in co-located collaboration with multiple input devices . *Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, 2008. New York, USA, pp. 1149-1152.
- NUI: Natural User Interface Group, <http://wiki.nuigroup.com/>, last retrieved February, 2011.
- Pinelle, David et al, The effects of co-present embodiments on awareness and collaboration in tabletop groupware, 2008. *Proceedings of graphics interface 2008*. Windsor, Canada, pp. 1-8.
- Surface: Welcome to Surface, <http://www.microsoft.com/surface/>, last retrieved February, 2011.
- Thaller, Georg Erwin, 2002. *Interface Design – Die Mensch-Maschine-Schnittstelle gestalten*; Software&Support Verlag, Frankfurt.

# INTERACTION DECONSTRUCTION METHOD FOR USABILITY REQUIREMENTS ENGINEERING

Lamminen Juha, Rousi Rebekah and Saariluoma Pertti

*Department Computer Science and Information Systems, University of Jyväskylä  
P.O. Box 35, FI-40014 University of Jyväskylä*

## ABSTRACT

This paper presents a novel usability requirements engineering method called Interaction Deconstruction, which supports user interface design in the interaction design process. The basic idea of the Interaction Deconstruction is to break down user actions into sub-actions and define, for each sub-action, the relevant usability-related attributes. This method is composed of three steps: defining the set of organized actions, constructing an organized tree structure, and setting the usability- and user experience related attributes for node states. The Interaction Deconstruction method differs favorably from existing usability requirements engineering methods. Firstly, it supports defining both usability and user experience requirements in parallel. Secondly, it utilizes psychological knowledge to engineer scientifically justifiable requirements. Thirdly, our method defines the relevant user-related attributes at the level of sub-actions. This is particularly beneficial in situations when interaction tasks are easy to separate into several sub-actions.

## KEYWORDS

Non-functional requirements, usability attributes, usability requirements, user experience requirements, requirements engineering

## 1. INTRODUCTION

Today it is essential to develop requirements engineering (RE) methods and techniques for human-technology usability and user experience design. This research is needed for industry to increase the speed of design and safeguard reliable outcomes and to improve communication between human and technical interaction usability design processes. Currently it is common to apply intuitive usability design processes in industry. Designers, who rely on scientific procedures when designing the electronics or mechanics of machines, tools and devices, may totally forget this attitude when they begin to construct interaction (Pahl et al, 2005). Nevertheless, there is no reason to assume that users are less complex than weather conditions or electronic circuits.

By usability and user experience design we refer to a design process that aims at designing interactive software products to support people in their everyday life. The process of interaction design typically involves four main activities (Preece et al, 1994): identifying needs and establishing requirements, developing alternative design, building interactive concepts and solutions, as well as evaluating the prototypes and solutions. In this paper we are interested in the process of establishing usability requirements for interactive software products.

The usability requirements definition process has mainly been discussed in usability engineering (UE) and user-centered-design (UCD) (e.g. ISO 9241-11; Mayhew, 1999; Nielsen, 1993; Preece et al, 1994; Quesenbery, 2003; Wixon and Wilson, 1997). However, the existing methods and techniques for usability requirements do not include support for facilitating the following issues:

- defining user experience requirements
- decomposing usability requirements into sub-requirements
- analyzing the usability-related attributes at the level of sub-actions
- clearly distinguishing usability and user experience components or cognitive and emotional usability
- utilizing scientific psychological knowledge

To solve the above problems and strengthen the definition of the usability requirements process, we propose a novel usability requirement analysis method called Interaction Deconstruction. This method is composed of three steps: defining a set of organized actions, which enable users to develop the technology (i.e., machine or program) towards the intended state, constructing an organized tree structure while keeping in mind the way in which users can reach their action goals, and setting usability-related attributes for node states.

We have strongly exploited earlier research on usability and user experience attributes to involve two of the most current and important interaction design aspects in our method: 1) how can people best use technology (e.g. traditional usability thinking)? And, 2) do people like to use a given technology (e.g. emotional usability or user experience research)? Our method borrows ideas from RE methods such as hierarchical task analysis in the experimental situation and analysis of user actions. Moreover, it goes further to emphasize the importance of usability and user experience attributes as the focus point of usability requirements engineering. Our method suggests analyzing user actions on the level of sub-actions (similar to the hierarchical task analysis) but our method also suggests adding the relevant usability-related attributes for each sub-action. Finally, our method supports the usability-related RE process of new and innovative user interfaces (UIs), overcoming many of the problems noted by literature when attempting to apply more traditional RE methods to similar tasks.

This paper is organized as follows: In section II a review of existing methods and techniques for usability requirements is presented; in section III, the definitions of traditional usability and user experience (in terms of related attributes) are presented; section IV presents the Interaction Deconstruction method; finally, in the last section, we summarize the results.

## 2. USABILITY REQUIREMENTS

The requirements can traditionally be divided into functional and non-functional requirements. The former can be seen as a system service or function (the intended actions of the system) and the latter as a constraint on the system or development process (how the system should perform these) (Kotonya and Sommerville, 1998). While the interaction design process concerns both functional and non-functional requirements (Preece et al, 1994), in this paper we are interested in usability requirements and user experience requirements which are non-functional.

Usability requirements are qualitative or quantitative usability goals for a computer system. Qualitative requirements can be subjective goals, which are not easy to measure (e.g. user satisfaction). Quantitative requirements are objective and measurable goals (e.g. the time taken to complete the task) (Stone et al, 2005). Qualitative usability goals can be seen as general goals that guide design, whereas quantitative usability goals can serve as acceptance criteria during usability evaluation (Mayhew, 1999). User experience requirements refer to the requirements concerning emotional aspects (e.g. using the intended system should be fun, enjoyable, and motivating). According to Preece (Preece et al, 1994) user experience requirements are typically subjective and difficult to quantify.

Usability attributes play a vital role in the usability requirements definition process. As an example, in the usability goal setting process by Wixon and Wilson (1997), the relevant usability attributes are determined based on specifying and categorizing the users and conducting task analysis. Also, Nielsen (1993) for example, has suggested that usability is associated with five usability attributes (learnability, efficiency, memorability, errors and satisfaction) and these attributes must be prioritized based on user analysis and tasks analysis when setting usability requirements. Nevertheless, there are still many open issues in developing a rational design process for usability.

Attributes turn the gaze towards evaluation rather than design. They are difficult to measure objectively and the outcome does not give us any idea about what should be done, when some objective outcome has been achieved. Knowledge of deficient learnability does not yet give any idea about what would be an ideal usability solution in that respect. Secondly, these traditional methods for usability requirements engineering (e.g. Mayhew, 1999; Wixon and Wilson, 1997) do not cover emotional and user experience requirements at all. Thirdly, few guidelines exist, or perhaps none at all, on how to define user experience requirements together with usability requirements and how to incorporate these two aspects of usability design into a rational design process. Here, we suggest that usability and user experience attributes and requirements be

considered together in the usability-related requirements engineering process. Fourthly, traditional usability requirements engineering methods do not give clear (if any) instructions on how to incorporate psychological knowledge (i.e., memorability, learnability etc.) into this process. This, in turn leads to intuitive (folk psychological understandings) adoption and application of psychological concepts.

Current usability requirements engineering literature indicates that there is an obvious lack of empirical research reports on quantitative usability requirements in practice (Kantola and Jokela, 2007), and that better methods for managing, organizing, and defining quantitative usability requirements should be developed (Jokela et al, 2006). Instead of quantitative usability requirements, literature presents many empirical research reports concerning qualitative usability requirements (e.g. Hermann and Heidmann, 2002; Tsalgatidou et al, 2003). However, these empirical research reports typically refer to requirements that focus on technological limitations such as strong bandwidth variability, unpredictable disconnections, limited computing power and small memory size. This approach leads to usability requirements stemmed from technology (e.g., possibility to offline operation), in which concrete user interface design details are considered at a very general level (e.g., "User interface should be very simple and user friendly and the amount of presented information content limited and well specified"). Of course, technology related aspects need to be noticed when usability requirements are defined, but we suggest that much more scientific psychological knowledge about humans as users is needed as well.

The goal of user interaction is to develop the technology towards an ideal end state. This is a state in which people can realize their action goals. A person willing to pay bills on the net wants to find a webpage which makes it possible. The task of usability engineering is to make reaching the end state via the technology easy and pleasant. However, to reach its goal, usability engineering should have clear processes. One important part of this work is to define the precise process of reaching the end state of technology, which often presupposes proceeding through a number of steps. Therefore, we suggest that the first step towards utilizing scientific knowledge in usability-related requirements engineering is to analyze interaction tasks on the level of sub-tasks. This is where the user is in the process of working towards the ideal technologically assisted state. Once these steps have been achieved, we can define the relevant usability and user experience attributes on the level of the defined sub-actions.

### 3. FROM USABILITY TO USER EXPERIENCE

In this section we present the definitions of traditional usability and user experience in terms of related attributes. This chapter is relevant for understanding the method of interaction deconstruction. This chapter is especially relevant in terms of understanding sub-section 4.3, where we explain how to set the usability-related attributes of node states. Further, this chapter provides a background of usability traditions, against which we apply our psychological approach to understanding the content of usability requirements. The chapter begins with describing the earlier views towards usability and then progresses into discussion of more contemporary approaches leading to user experience. We use the concepts of fun and enjoyable in sub-sections 3.1 and 3.2 to illustrate what we mean by psychological content of user experience.

Shackel's (1991) definition of usability is one of the best-known. According to him, the attributes of usability are effectiveness, learnability, flexibility and attitude. Effectiveness refers to performance in the accomplishment of tasks (e.g. speed and errors). Learnability is the degree of learning to accomplish a task (e.g. time to learn and retention.). Flexibility means adaptation to variations in tasks and/or environments, and attitude refers to user satisfaction with the system.

Nielsen (1993) suggests that usability is associated with five usability attributes: learnability, efficiency, memorability, errors, and satisfaction. By learnability, Nielsen refers to how well a novice user can use the system, while efficiency has to do with efficient use by an expert. Memorability stands for the ability to use the system occasionally, without the need to re-learn it every time it is used. As far as errors are concerned, the system should have a low error rate, and this attribute also includes users recovering from possible errors. Satisfaction refers to the requirement that the system should be pleasant to use. According to Nielsen (1993), these attributes are not unambiguous and they can, to a certain degree, conflict with one another. Here, Nielsen mentions satisfaction as being a key component of usability, however, as with much user experience scholarship that we will shortly mention, this does not give a clear definition of what statements such as 'pleasant to use' actually entail.



The definition of usability most often referred to is probably the one in ISO 9241-11 (1998) which defines usability as consisting of three attributes: effectiveness, efficiency, and satisfaction. Effectiveness refers to accuracy and completeness in achieving specified goals. Efficiency means resources expended in relation to the accuracy and completeness in achieving goals. Satisfaction refers to freedom from discomfort, and positive attitudes towards using the product. Again, this definition which is often repeated when referring to usability and user experience points towards feelings experienced during satisfaction. However, the way that they exist within this definition is still on an extremely general level. We are still left wondering what the core components of discomfort and positive attitudes actually are.

ISO 9126 (2001) identifies usability as one of six different software quality categories that are relevant during product development. The other categories are functionality, reliability, efficiency, maintainability, and portability. Furthermore, the usability category is divided into five attributes: understandability, learnability, operability, attractiveness, and usability compliance. Understandability refers to the capability of a software product to enable the user to understand whether the software product is suitable or not. Learnability is the capability of a software product to enable the user to learn its application. Operability means the capability of a software product to enable the user to operate and control it. Attractiveness refers to the capability of a software product to attract the user. Usability compliance refers to compliance with published style guides or conventions for user interfaces. Here, the aspect of attractiveness is raised, but it seems to be more vague than the aspect of satisfaction, due to the abstractness of what is actually attractive within the product – to whom is it attractive, and for what reasons?

We will return back to the 1990s in a moment, but jumping forwards a decade, Quesenbery (2003) suggests that the user dimensions of usability are associated with five attributes (five Es) described by the adjectives effective, efficient, engaging, error-tolerant, and easy-to-learn. It can be seen that part of the main motivation behind this classification is that computer systems have moved beyond being just work-based (Stone et al, 2005). Although the five Es are mainly based upon the ISO 9241-11 (1998), one different kind of attribute, called engaging, is suggested. Engaging refers to the degree to which the tone and style of the interface make the product pleasant or satisfying to use.

Moving back into the 1990s, and somewhat addressing this notion of engaging, Preece et al. (1994) argue that when designing interactive products, both usability and user experience are needed. The former refers to the following attributes: efficiency (efficient to use), effectiveness (effective to use), safety (safe to use), utility, learnability (easy to learn), and memorability (easy to remember how to use). The latter refers to fun, emotional fulfillment, rewarding experiences, support for creativity, aesthetically pleasing feelings, motivation, helpful aspects, entertainment, enjoyment, and satisfaction. From a psychological perspective, the following sub-sections elaborate on what is fun and enjoyable in relation to user experience. These are not of course the only components of user experience, others such as motivation exist, but we use these to demonstrate the psychological dimension.

### 3.1 Fun

Read et al. (2002) suggest that the relationship between fun and usability is associated with three dimensions: expectations, engagement, and endurability (especially in the context of measuring children's fun). Expectations refer to the expectation of use and subsequent perception. This means that fun is composed of two components, (a) the fun that is attached to the current event, and (b) the fun that is attached to the prior expectations of the user. As an example, if we have high expectations and an event is then perceived to be dull, we feel disappointed, and vice versa. Engagement refers to the positive and negative instantiations (e.g. smiling, laughing, concentration etc.). Endurability is composed of two facets: remembrance and accessibility. The former is based on the Pollyanna principle: the likelihood to remember things that we have enjoyed. The latter refers to the desire to perform a fun activity again. Here, it is important to remember that numerous factors affect what is perceived as fun. Most notably, people's notions of what is fun changes with age and for example experiences. Also, perceptions of fun alter depending on culture, social conditions, profession, education and so forth. Thus, careful attention needs to be placed during research towards these influential factors.

### 3.2 Enjoyable

Lin and Gregor (2006) have reviewed relevant studies of enjoyment (especially as it is experienced by users of the Web), and they suggest that the concept of enjoyment is necessarily characterized by engagement in an activity, the resultant positive affect, and the fulfillment of some need or desire. Engagement in an activity refers to a situation in which attention is deeply focused on some activity. The resultant of positive affect refers to positive emotions (e.g. pleasure, happiness, contentment etc.) which result from use experience. The fulfillment of some need or desire refers to the fulfillment of user needs, including unexpected needs.

## 4. INTERACTION DECONSTRUCTION METHOD

The main goal of the Interaction Deconstruction method is to develop a human-technology interaction model of relevant user interface actions which would allow users to use technology in the best possible manner. In other words, the users would be able to effectively reach their action goals through a user interface. The Interaction Deconstruction method is composed of three steps: defining the set of organized actions, constructing an organized tree structure, and setting usability-related attributes for node states (see Figure 1). In step 1, we define the relevant actions on the level of high-level goals. In this case we use the example of internet banking, and within this first step the problem of reaching this high-level goal is difficulty in paying a bill. In step 2, we move towards examining the low-level goals, by looking at the sub-tasks of the operational actions or interactive intentions. In step 3, we articulate the usability-related attributes. This method can be used for example to develop totally new user interfaces and replace/update existing user interfaces.

### 4.1 Defining the Set of Organized Actions

The first step of the Interaction Deconstruction method is to define the set of organized actions. In case a totally new and innovative user interface is developed, we suggest utilizing an intention-driven approach (Leppänen et al, 2010). This approach combines technological, social, and business viewpoints aiming to define the intention driven “use of the product” goals and requirements. In the Interaction Deconstruction method, these goals and requirements refer to the step defining the set of organized actions.

In the instance of replacing or updating the existing user interface, the set of organized actions refers to the certain user interface actions which are proposed to be repaired or replaced. For example, if a netbank system is needed to be renewed, the set of organized actions might refer to actions such as paying a bill, transferring money, watching stock markets, and buying pension insurance. There are several suitable methods and techniques for gathering and analyzing the relevant interactions from existing products (see e.g. Hoffman and Millitello, 2008).

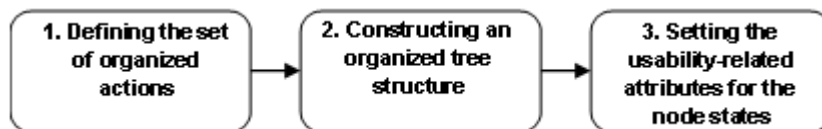


Figure 1. The steps of the Interaction Deconstruction method

### 4.2 Constructing an Organized Tree Structure

In step 2, user interaction is divided into a form of hierarchical tree, composed of node states and operations. At this point, one can use hierarchical task analysis (Annett, 2003), the goal of which is to break a task into subtasks and then into sub-subtasks if needed. As an example, if the set of organized actions refers to user interaction with a netbank system (the previous example), the certain action “paying a bill by using a netbank system” can be broken down into subtasks: 1.1 log in; 1.2 select new payment; 1.3 enter the required payment information (e.g. account number, beneficiary’s name, amount, reference number, and message);

1.4 confirm payment by using confirmation code; and 1.5 log out. This refers to the interactive intention or low-level goals of the user.

### 4.3 Setting the Usability-related Attributes for the Node States

Clearly, whether the user interface is a graphical user interface (GUI), a Web user interface (WUI) or a touch user interface, the user controls the states of interaction with the product by making selections (e.g. pressing the log in button). There are always node states and operations which move the system from one node to another. The task of the user is to make the decisions which give meaning or sense to the transitions from one state to another. As the nodes (e.g. 1.3 enter the required payment information) are contact points in interaction, we have to consider usability and user experience attributes at these points. As an example, when the user is entering the relevant payment information, preventing errors (usability attribute) is a more important attribute than fun or entertainment (user experience attribute). Furthermore, when the user is logging out, the feeling of trust with other user experience attributes are more important than usability attributes such as efficiency and effectiveness. It is important to find a balance between usability and user experience attributes on the level of sub-actions. This is the process which differs from previous usability requirement engineering processes.

Step 3 means that the nodes of the tree structure are equipped with appropriate traditional usability and user experience attributes. The traditional usability attributes (e.g. effectiveness, efficiency, memorability, and learnability) refer to factors which foster the capability of the user to make the correct decision in the node. They help people pick up relevant information for making the right decision. The user experience attributes (e.g. enjoyable, fun) give information about how user interface designers can implement interfaces which the users like to use. Both usability and user experience aspects are important when the user interface interaction model is built in action-oriented user interface design.

As an example, a radio button can be associated with a single action. The designer must understand how to differentiate the critical button from other buttons and how to communicate that this button is important in reaching the defined action state. On the other hand, the designer must also think about what the design should be like, in order to draw the appropriate experiential reaction. For example, should the users find the UI cool or prestigious? What is the intended experiential message that the owners of the program hope to transfer to the user? Thus, when considering user experience, the UI is considered from a different point of view. While the usability concepts help in making a device easy to use, the user experience concepts convey notions of what looks fine and is enjoyable. In this way, the same node should be analyzed from different perspectives. This leads into the construction of an organized tree structure description of the relevant interaction tasks. The hierarchical tree structure description is composed of node states (or sub-actions) with relevant usability and user experience attributes, plus operations between the nodes.

In addition to the obvious usability factors, we can analyze the above chain of sub-actions from the user experience perspective, starting from the netbank welcome and log in page. To ensure that a user will feel comfortable with entering their account number and password on the log in page, the user should be able to not only identify that the current and subsequent pages are secure via symbols such as the padlock in the bottom right-hand corner of the page, or various trusted page symbols, but also authenticity within the layout design itself. Visually, the initial page needs to conform to the schematic themes seen throughout the organization. It needs to express order (Lynch, 2010; Tractinsky, 2004; Tractinsky et al, 2006).

Once the user has logged in, as indicated by the nodes chart above, steps to be taken in order to reach the intended state need to be clear, concise and direct. Coinciding with the usability factor is the user experience of feeling that one understands the terminology of the steps, thus, simple language is necessary. The sense of control is then heightened by the interactive, personalizing attribute of the message function. When the steps have been completed and the user logs out, as mentioned before, it is necessary for maintaining the experience of trust, to include a message that the user is safely logged out. Thus, as within a narrative, there is closure to the set of actions.

This example illustrates how our understanding of user experience differs and expands upon the previous user experience models mentioned earlier in the article. Both Preece's and Quesenbery's models combine elements of usability (Preece – helpful; Quesenbery – effective, efficient, error tolerant and easy to learn), with elements encouraging active positive emotions (the rest of Preece's list and 'engaging' from Quesenbery's list). The problem with these elements is that they are not related to the experiential nature of

banking. Perhaps in earlier times when people visited bank branches in person, we were attracted by the friendly, helpful service which covered a number of Preece's listed elements. But when dealing with an online UI responsible for the management of one's financial well-being, the experience the user is hoping to gain covers another side of the cognitive-emotional spectrum: the user's sense of trust and security is heightened by the feeling of control induced by a clear, understandable and straight-forward user interface operating on both visual/aesthetic and cognitive levels; and trust in the UI itself also comes through consistency (authenticity) of layout and symbols, operating on an aesthetic level. Rousi et al. (2011) discuss this in terms of the dimensions of user experience, which divides mental contents into four different (yet somewhat overlapping) categories: emotional, aesthetic, cognitive and practical.

## 5. CONCLUSION

Usability-related attributes play a major role in the Interaction Deconstruction method. In this method it is essential to consider usability and user experience attributes not only on the level of a single interaction task but also on the level of single nodes. In this way, it is also possible to define usability-related requirements on the level of single state nodes. Consequently, this method is particularly suitable for UI-intensive systems with high demands to reduce risks.

Usability and user experience concepts are interrelated. If the users are not able to reach their goals with technologies in an easy manner, they become irritated. They do not like slow or incomprehensible systems. It is also possible that a beautiful and enjoyable interface design makes it difficult to find and comprehend crucial information. In this way, the two sets of attributes are often intimately connected to each other. Our action-oriented tree structure analysis gives roles for both types of usability-related attributes.

Our goal was to outline a design approach which is based on the idea of commencing interface design with the analysis of human actions and separately investigating the issues of how users can use interfaces and how they like to use them. The main motivation is that we have to unify "can" and "like" –perspectives in practical interaction design. The action-oriented user interface design and development scheme is composed of three steps: defining a set of organized actions, constructing an organized tree structure, and setting usability-related attributes for node states. Interaction Deconstruction defines the relations between usability and user experience attributes. It enables us to understand why the users really use the technology, how they could best deploy it and finally, what makes them feel good when they reach their action goals through the user interface.

In this model, step 3, setting the usability-related attributes for the node states, is considered central as it differentiates the Interaction Deconstruction method from other usability requirements engineering methods. Step 3 defines the user experience requirements in relation to usability. This is achieved by clearly distinguishing usability and user experience components through charting nodes of interaction. Through doing this, scientific psychological knowledge is applied to generate a more in-depth understanding of user experience.

## REFERENCES

- Annett, J., 2003. *Hierarchical Task Analysis*. In D. Diaper & N. Stanton (Eds.), *The Handbook of Task Analysis for Human-Computer Interaction* (pp. 67–82). Lawrence Erlbaum Associates, Mahwah, NJ, USA .
- Hermann, F. and Heidmann, F., 2002. User requirement analysis and interface conception for a mobile, Location- Based Fair Guide. *Proceedings of the 4th International Symposium on Human-Computer Interaction*. pp. 388-392.
- Hoffman, R. and Millitello, L., 2008. *Perspectives on cognitive task analysis: Historical Origins and Modern Communities of Practice*. Taylor and Francis, New York.
- Jokela, T. et al, 2006. Quantitative Usability Requirements in the Development of the User Interface of a Mobile Phone. A Case Study. *In Personal and Ubiquitous Computing* Vol. 10, No. 6, pp. 345-355.
- International Organization for Standardization [ISO], 1998, March. *Ergonomic requirements for office work with visual display terminals (VDTs), Part 11: Guidance on usability*. (Standard No. 9241-11). Geneva, Switzerland: ISO.
- International Organization for Standardization [ISO] and International Electrotechnical Commission [IEC], 2001, June. *Software engineering: Product quality–Quality model*. (Standards No. 9126-1). Geneva, Switzerland: ISO.

- Kantola, N. and Jokela, T., 2007. Determining high level quantitative usability requirements: a case study. *Proceedings of the 12th international conference on Human-computer interaction: interaction design and usability (HCI'07)*, Julie A. Jacko (Ed.). Springer-Verlag, Berlin, Heidelberg, pp. 536-543.
- Kotonya, G. and Sommerville, I., 1998. *Requirements Engineering, Processes and Techniques*. John Wiley & Sons, Chichester, England.
- Leppänen, M. et al, 2010. A framework for intention-driven requirements engineering of innovative software products. *In: 19th International Conference on Information Systems Development (ISD 2010)*, Prague, Czech Republic, August 25 - 27 (in print).
- Lin, C.H. and Gregor, S., 2006. Understanding and assessing the extent of enjoyment of Web experiences. *Proceedings of the 17th Australasian Conference on Information Systems (ACIS 2006)*, Adelaide, Australia.
- Lynch, P., 2010. Aesthetics and Trust: Visual Decisions About Web Pages. *Proceedings of AVI'10*, Rome, Italy.
- Mayhew, D. J., 1999. *The usability engineering lifecycle: A practitioner's handbook for user interface design*. Morgan Kaufmann Publishers, San Francisco.
- Nielsen, J., 1993. *Usability engineering*. Academic Press, Boston.
- Pahl, G. et al, 2005. *Engineering design*. Springer, Berlin.
- Preece, J. et al, 1994. *Human-computer interaction*. Addison-Wesley, Wokingham, England.
- Quesenbery, W., 2003. *The five dimensions of usability*. In M.J. Albers, and B. Mazur (Eds.) *Content and Complexity: Information Design in Technical Communication* (pp. 81–102). Lawrence Erlbaum Associates, Mahwah, NJ, USA.
- Read, J.C. et al, 2002. Endurability, engagement and expectations: Measuring children's fun. *Proceedings of Interaction Design and Children*, International workshop, Shaker Publishing, Eindhoven.
- Rousi, R. et al, 2011. Unpacking the Contents: a conceptual model for understanding user experience in user psychology. *Proceedings of the Fourth International Conference on Advances in Computer-Human Interactions February 23-28, 2011*, Guadeloupe, France.
- Shackel, B., 1991. *Usability: Context, framework, definition, design and evaluation*. In B. Shackel & S. Richardson (Eds.), *Human factors for informatics usability* (pp. 21–38). Cambridge University Press, Cambridge, MA, USA.
- Stone, D. et al, 2005. *User interface design and evaluation*. Morgan Kaufmann Publishers, San Francisco.
- Tractinsky, N., 2004. Using Ratings and Response Latencies to Evaluate the Consistency of Immediate Aesthetic Perceptions of Web Pages. *Proceedings of the Third Annual Workshop on HCI Research in MIS*, Washington, D.C.
- Tractinsky, N. et al, 2006. Evaluating the Consistency of Immediate Aesthetic Perceptions of Web Pages. *International Journal of Human- Computer Studies*, Vol. 64, No.11, pp 1071–1083.
- Tsalgatidou, A. et al, 2003. Mobile E-Commerce and Location-Based Services: Technology and Requirements. *Proceedings of the 9th Scandinavian Research Conf. on Geographical Information Science*, Espoo, FIN, pp. 1-14.
- Wixon, D. and Wilson, C., 1997. *The usability engineering framework for product design and evaluation*. In M. Helander, T. Landauer, & P. Prabhu (Eds.) *Handbook of human-computer interaction* (pp. 653–688). Elsevier, Amsterdam.

# COGNITIVE CHUNKS EXTRACTED FROM EYE-TRACKING RECORDS OF WEB PAGE READERS, USING ISOLATED SACCADDES AS A DELIMITER

Noriyuki Matsuda<sup>1</sup> and Haruhiko Takeuchi<sup>2</sup>

<sup>1</sup>*Department of Systems and Information Engineering, University of Tsukuba  
1-1-1 Tennou-dai, Tsukuba, 305-8573, Japan*

<sup>2</sup>*National Institute of Advanced Industrial Science and Technology (AIST)  
Tsukuba Central 6, 1-1-1 Higashi, Tsukuba, Ibaraki, 305-8566, Japan*

## ABSTRACT

Few previous eye-tracking studies incorporated the psychological notion of chunking, a meaningful cognitive unit of information. In the present work, we constructed fixation sequence lists nesting chunks, using isolated saccades as a delimiter. Chunks were extracted from the time-stamped records of the fixation sequences that were coded according to the 5x5 segments imposed on the display. The overwhelming majority chunks consisted of one or two fixations. Most within-chunk distances were zero or one, while the between-chunk distance was relatively dispersed with the modal distance at one, followed by zero. There was good agreement between the rankings of between- and within-chunk loops among the primary and secondary segments on the total number of loops. We found some indications about the layout effect, possible attributable to the presence of sub-area on the right most segments.

## KEYWORDS

Eye-tracking, fixations, saccades, chunking, loops

## 1. INTRODUCTION

When people view a web page, they move their eyes to gain useful or interesting information. At times, their eyes stay in fairly limited space, and at other times the eyes rapidly move to a new, remote target. Eye-tracking researchers called them fixations and saccades, respectively. It is generally held that people engage in the analytical task on the focused aspects of the stimulus during fixations, whereas substantive cognition is unlikely to occur in saccades (see, Fulton, 2009; Lorigo et al., 2008). Since reading texts or observing scenes involves a series of fixations not entirely in succession but with occasional pauses, it seems that cognition of viewers proceeds by chunking information instead of seamless flows. Actually, the punctuation in writing has evolved in essence to assist and/or force such chunking (see, Parkes, 1992). The present paper introduces a new approach to eye-tracking data by using chunking as a key factor.

Frequent eye movements are inherent in our perceptions, since they help us bring items of interest to the fovea where visual acuity is best (Melcher, 2008). That is, our eyes are rarely fixed on a single point over time. Hence, the notion of an eye-fixation actually pertains to the gaze-points clustered in a limited space over certain duration sufficient for visual processing, for instance, within 1° to 2° of visual angle and over 100-200 msec (Goldberg et al., 1999; Salvucci et al., 2000).

Put slightly more technically, a fixation is determined by the consecutive gaze points that satisfy certain spatial and temporal criteria. Spatial dispersion of points within a fixation results from short, micro-saccades as opposed to longer, macro-saccades by which eyes land on a spot outside of the preceding fixation sphere. A new fixation may ensue a macro-saccadic move if not followed by further macro-saccade(s). Chunks in general consist of one or more fixations. Figure 1 illustrates successive fixations without (left) and with (right) an isolated point not participating in any fixation. In the present paper, isolated point is treated as a delimiting marker of chunking. An analogy seems to hold, in a rough sense, between the <gaze-point, fixation, chunking> relationship and the <letter, word, phrase> relationship.

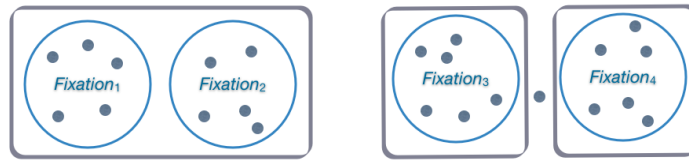


Figure 1. Two consecutive fixations within a chunk (left) and in separate chunks (right)

Like the conventional heatmaps (see Granka et al., 2006; Josephson, 2004; Pan et al., 2004; Matsuda et al., 2011a, b), scanpaths (Goldberg et al., 1999; Nielsen et al., 2010) and network analysis (Matsuda et al., 2009, 2010, 2011a, b, c), we construct fixation sequences from the raw data of gaze-points. The difference is that the new sequences nest fixation chunks in them as opposed to the flat structure in the conventional approaches; e.g.,

Nested structure:  $[ [F_1 F_2], [F_3], [F_4], \dots, [F_i F_{i+1} F_{i+2} \dots], \dots ]$  /\*  $F_i$  denotes the  $i$ th fixation.  
 Flat structure:  $[ F_1 F_2 F_3 F_4 \dots F_i F_{i+1} F_{i+2} \dots ]$

The advantage of nested structure seems to be clear by virtue of embracing two kinds of attentional shifts, i.e., within- and between-chunk shifts to the extent, saccades are preceded by an attentional shift to the target position (see Goldberg et al, 1999; Hayhoe, 2004).

When fixations are coded in terms of areas-of-interest (AOI) or region of interest (ROI), an attentional shift may produce landing in the same area, known as a looped transition in the network terminology. Such a loop may be regarded as a reflection of sustained attention (Matsuda et al. 2011b, c) in the flat structure of fixations. Now that the nested structure is in hand, we are able to make a finer distinction about loops--within- and between-chunk loops. The former reflects a genuine sustained attention, whereas the latter actually represents immediately recurrent attention.

It would be ideal if we could relate chunking patterns to cognitive processes at high levels. Until all the substantive issues (see Lorigo, 2008; Martinez-Conde, 2004; Rayner, 1998) and methodological ones (see van der Lans et al. 2010; Poole et al., 2005) are solved, we should keep paving the road by developing analytical tools. In the present work, the tool is the extraction of chunks as explained below.

gaze pint	coordinate	time
1	$x_1, y_1$	$t_1$
2	$x_2, y_2$	$t_2$
3	$x_3, y_3$	$t_3$
4	$x_4, y_4$	$t_4$
5	$x_5, y_5$	$t_5$
6	$x_6, y_6$	$t_6$
7	$x_7, y_7$	$t_7$
...	...	...

gaze pint	coordinate	time
1	$x_1, y_1$	$t_1$
2	$x_2, y_2$	$t_2$
3	$x_3, y_3$	$t_3$
4	$x_4, y_4$	$t_4$
5	$x_5, y_5$	$t_5$
6	$x_6, y_6$	$t_6$
7	$x_7, y_7$	$t_7$
...	...	...

Fixation-identification algorithm, given the 50Hz sampling rate.  
 Step 0 Put the gaze-points in the scan list (SL)  
 Step 1 Copy the first five points from SL to the working list (WL)  
 Step 2 If the point in WL all falls within the circle  
     Repeat adding a new point in SL to WL  
     until additional point falls outside of the circle  
     Register the centroid of the points in WL as a new fixation  
     Remove from SL the points that were copied or added in WL  
 Otherwise,  
     Remove the first point in SL  
 Step 3 If there are more points in SL  
     Go to step 1.

Figure 2. A scanning window that shifts its position (left) and expands the coverage (right)

Figure 3. The procedure for fixation identification

Given our stance, let us explain the procedures for extracting chunks that are closely related to the identification of fixations from gaze points. According to the dispersion model among various alternatives (Salvucci et al., 2000), a fixation is defined as the centroid of the gaze points falling within a certain radius over certain duration. Operationally, a window is applied to scan the list of sequential gaze points to find such clusters of points. The window moves along the list or expands to include additional points as shown in Figure 2.

With the sampling rate for instance at 50Hz, we obtain five observations in every 100 msec. If these points fall within a given radius, we keep adding a new point to the window until a new point falls outside of the circle. When it comes to a halt, we register the centroid of the points in the window as a fixation point. After renewing the window, repeat this process from the beginning as long as there are more points remaining in the data. The algorithm shown in Figure 3 and the pictorial presentation in Figure 4 will help readers understand the process.

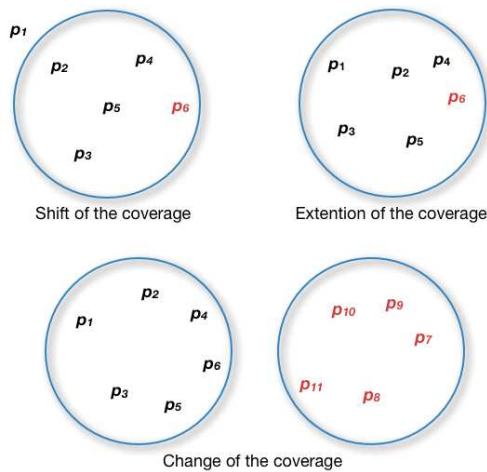


Figure 4. Basic patterns of the window coverage

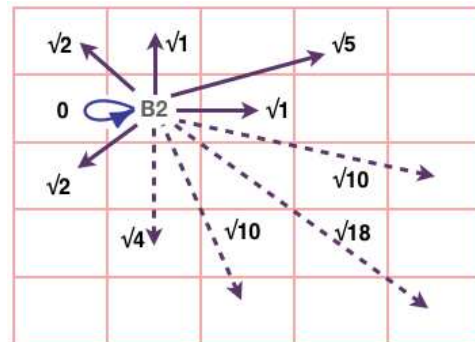


Figure 5. Vector length between segments

By recording fixations with time-stamps, we are able to extract chunks from a fixation sequence, using isolated points as a delimiter (see Figure 1). To be more precise, the presence of such point(s) causes at least 20 msec time-lag between two successive fixations, i.e., the last point of fixation<sub>k</sub> and the first point of fixation<sub>k+1</sub>. That means, it is the time elapsed between fixations that actually delimits chunks.

Before concluding this section, a note seems in order about the distance between fixations illustrated in Figure 5. A fixation located in a segment may have a successor landing again in the same segment (i.e., loop) or in another segment. The distance between the fixations will be approximated by the vector length between segments, by assigning one to each horizontal and vertical move. Alternatively, one can compute Euclidian distance from the centroids of the fixations if exact measurement is desired. However, it makes treatment of loops difficult. In the present study, we opt for conceptual ease of interpretations.

## 2. METHOD

**Subjects (Ss)**--Twenty residents (7 males and 13 females) living near a research institute AIST, Japan, participated in the experiment. They had normal or corrected vision, and their ages ranged from 19 to 48 years (30 on the average). Ten Ss were university students, five were housewives, and the rest were part-time job holders.

**Stimuli**--The top pages of ten commercial web sites were selected from various business areas, including airline, e-commerce, finance and banking. The pages were classified into Types A, B and C depending on the layout of the principal part beneath the top layer (See Figure 6). The main area of Type A was sandwiched between subareas, while the main areas of Types B and C were accompanied by a single sub-area either on the left (B) or the right (C).

Type A			Type B		Type C	
Top layer			Top layer		Top layer	
Sub	Main	Sub	Sub	Main	Main	Sub

Figure 6. Layout types

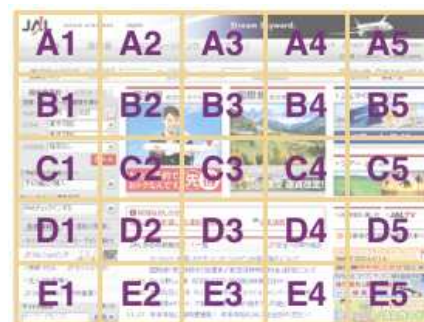


Figure 7. Segment coding



**Apparatus and procedure**--The stimuli were presented with 1024 × 768 pixel resolution on a TFT 17" display of a Tobii 1750 eye-tracking system at the rate of 50Hz. The web pages were randomly displayed to the Ss one at a time, each time for the duration of 20 sec. The Ss were asked to browse each page at their own pace. The English translation of the instructions is as follows: "Various Web pages will be shown on the computer display in turn. Please look at each page as you usually do until the screen darkens. Then, click the mouse button when you are ready to proceed." The Ss were informed that the experiment would last for approximately five minutes.

**Segment coding** --A 5x5 mesh was imposed on the effective part of the screen stripped of white margins which had no text or graphics. The segments were sequentially coded as shown in Figure 7 with a combination of the combination of alphabetical and numerical labels for rows and columns, respectively: A1, A2, ..., A5 for the first row; B1, ..., B5 for the second; ...; and, E1, ..., E5 for the fifth.

**Fixation sequences nesting chunks**--The raw tracking data for each subject comprised xy-coordinates of the gaze-points with time-stamps and were transformed to the fixation points under the condition that the subject's eyes stayed within a radius of 30 pixels for at least 100-msec. The fixations points were computed from the centroid of the graze points clustered in the given radius. Then, each fixation record was translated into code sequences according to the segments in which the fixation points fell. Each of these sequences was further transformed to a sequence nesting chunks consisted of a single fixation or successive fixations with no time-lag exceeding 20 msec.

### 3. RESULTS

Among the 10 top pages used as stimuli, Page 5 was eliminated due to the broad white space. Page 4 was also eliminated because of anomalies in the data. The resulting pages, referred to as TP $n$  hereinafter, were subjected to the analysis:

TP1, 3, 6 and 8 (Type A); TP2 and 9 (Type B); TP7 and 10 (Type C)

Type label will be prefixed to TP $n$  like A\_TP1 to save space in this section and Discussion.

The following example of a nested sequence will help understand special terminologies:

[ [B2B2] , [B2] , [A1] , [B4B4B3] , [B3]... ]

Where [B2] , [A1] and [B3] are singletons, and [B2B2] is a homogenous chunk composed of a loop, while [B4B4B3] is heterogeneous containing a loop and another fixation. Two B3's in successive chunks constitute between-chunk loop.

#### 3.1 Length and Patterns of Chunks

The simplest, or shortest, chunks are singletons, containing only one fixation. The singleton chunks outnumbered longer ones in the proportion of occurrence as listed in Table 1. The majority chunks ( $\geq .533$ ) were of length one. The second most frequent ones were of length two falling within the range of [.200, .27.5]. Put

Table 1. The proportions of chunks by length (%)

Length	Type A				Type B		Type C	
	TP1	TP3	TP6	TP8	TP2	TP9	TP7	TP10
1	60.4	58.4	58.9	53.3	55.1	54.1	57.5	56.5
2	21.9	22.1	22.6	25.1	26.4	27.5	20.0	20.3
3	9.8	8.9	6.9	12.2	7.0	11.1	9.6	10.0
$\geq 4$	7.9	10.6	11.6	9.5	11.5	7.3	12.9	13.3
Total (N)	100.0 (593)	100.0 (539)	100.0 (552)	100.0 (550)	100.0 (573)	100.0 (615)	100.0 (544)	100.0 (572)

together, the chunks of length one or two accounted for more than 76.7% (C\_TP10) to 82.3% (A\_TP1) of the observed chunks.

Table 2. The level of realization of different patterns by length

Length	Type A				Type B		Type C	
	TP1	TP3	TP6	TP8	TP2	TP9	TP7	TP10
1	1.00	1.00	.960	1.00	1.00	1.00	1.00	1.00
2	.118	.083	.114	.106	.104	.112	.086	.080
3	.003	.002	.002	.003	.002	.003	.003	.003

Concerning to the patterns of chunks, the singletons were the extremely diversified: All of the 25 segments appeared as singletons with a sole exception of E3 on A\_TP6. In other words, all or almost all possible patters were realized in singletons.

Shown in Table 2 are the proportions of the realized patterns among maximally possible ones, by chunk length. The base was computed by  $25^{length}$ , permitting repeated use of codes. Clearly, the realized patterns of the non-singleton chunks were much less diversified than singletons as a whole.

Nevertheless, homogenous chunks of length two (i.e., loops) quite varied in constituents. As shown in Table 3, on the TPs of Types A and B, the loops were found in 80 to 92% of the 25 segments, while on the TPs of Type C was somewhat smaller (68%). Besides, the homogenous chunks tended to occur more frequently than heterogeneous ones. At least five most frequent chunks were homogenous across TPs.

For the chunks of length three, the homogeneity was limited only to the complete ones and those with partial loops such as A1A1B2 or B2A1A1 were treated as heterogeneous. In consideration of this strict criterion, the diversity among homogenous chunks was noteworthy, ranging from 24% to 44%. Also, at least the most frequent chunks were homogenous across TPs except for B\_TP2. Within-chunk loops will be examined later.

Table 3. The proportions of the segments participated in homogenous chunks (%)

Length	Type A				Type B		Type C	
	TP1	TP3	TP6	TP8	TP2	TP9	TP7	TP10
2	84.0	80.0	84.0	92.0	92.0	84.0	68.0	68.0
3	44.0	24.0	24.0	36.0	28.0	40.0	28.0	40.0

### 3.2 Between- and within-Chunk Distances

The between-chunk distance is the length of a vector (Figure 5) connecting the last fixation of the preceding chunk to the first fixation of the successor, whereas the within-chunk distance is the sum of the vector lengths between successive fixations within a given chunk.

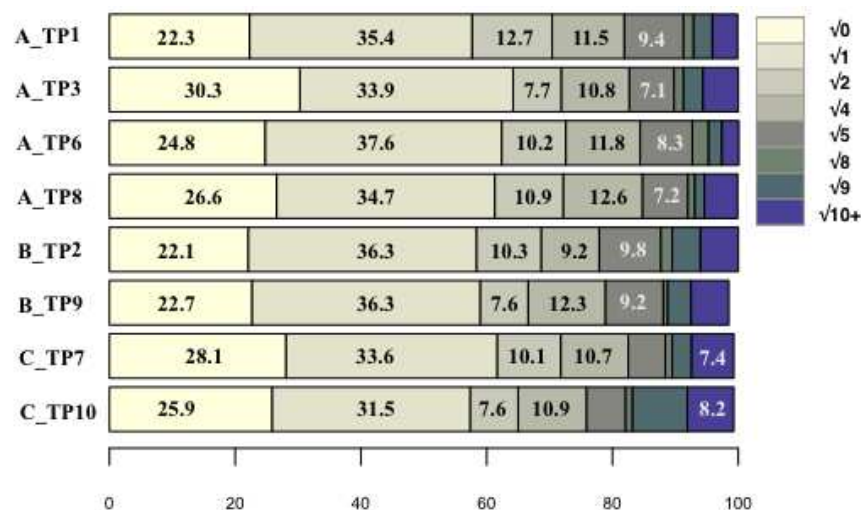


Figure 8. Proportions of the between-chunk distance

Note: The percentage scores are shown for the five largest proportions.

As shown in Figure 8, the modal between-chunk distance was  $\sqrt{1}$  across TPs, resulted from the shift to either the horizontal or vertical adjacent segment. The second modal distance was zero, resulted from the recurrent loop that was intervened by one or more isolated saccades. The scores of the five largest proportions are presented in Figure 8. Their cumulative values ranged from 85.1 (C\_TP7) to 92.7% (A\_TP6). Those largest values were found within the distance of  $\sqrt{5}$  on the TPs of Types A and B, the TPs of Type C yielded non-trivial proportions of the long distance moves ( $\sqrt{10}$  or more). Type C also differed from Types A and B with respect to the diversity of homogeneous chunks of length two.

Of interest is the lack of consistent prevalence of distance  $\sqrt{4}$  in comparison with longer ones ( $\sqrt{5}$  or more). It was the third modal distance only on A\_TP1 and B\_TP2. On the rest TPs, the move in distance  $\sqrt{5}$  was more prevalent.

In contrast to the between-chunk distance, the within-chunk distance was predominantly resulted from immediate loops followed by the shifts to the horizontally or vertically adjacent segments (see Figure 9). The

fact that the loops were modal is congruent with the results of chunks of length two and three, although the Figure 9 reflects the patterns of the chunks composed of three or more fixations. The cumulative proportions of the first two primary distances ranged from 89.8 (A\_TP8) to 95.4% (C\_TP10).

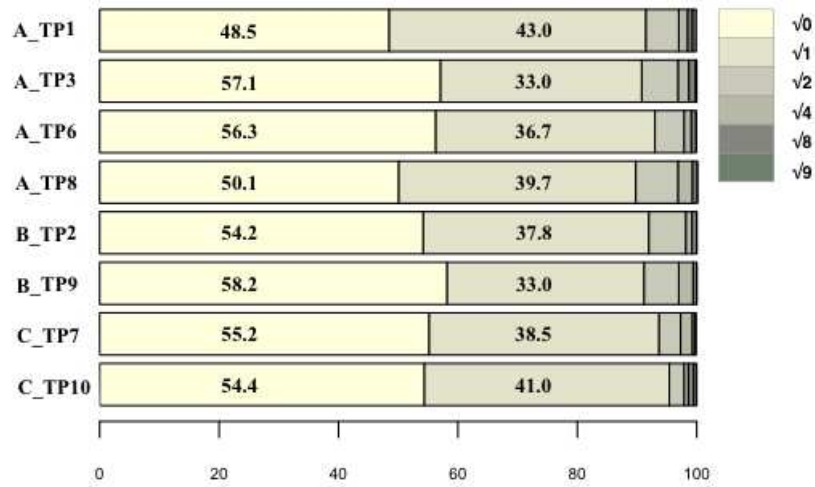


Figure 9. Proportions of the within-chunk distance  
 Note: The percentage scores are shown for the two largest proportions.

### 3.3 Between- and within-Chunk Loops

Sufficient number of total loops (*NLop*) was needed to examine both between- and within-chunk loops (to be referred to as *BLoP* and *WLoP*, respectively, hereinafter). Thus, we selected the segments with *NLop* equal to or above the median for each TP. The rank-order correlation coefficients (*Kendall's*  $\tau$ ) between *NLop* and *WLoP* were high across TPs ( $\geq .782$ ), whereas those between *NLop* and *BLoP* varied from .394 to .82.

Type A						Type B						Type C											
TP1	1	2	3	4	5	TP3	1	2	3	4	5	TP2	1	2	3	4	5	TP7	1	2	3	4	5
A					6 9 8	A	1 1 1					A	1 2 1				6 6 5	A	2 2 1				5 1 2 5
B	2 5 1	1 1 2			4 3 5	B	2 2 2		4 3 5			B	3 4 3		2 1 3	4 7 2		B	4 4 4	6 6 7			3 3 3
C				6 1 2 5		C	6 8 6				5 7 3	C		7 5 7				C					
D			5 9 4		3 2 3	D	3 4 3		7 5 8			D				5 3 6		D	7 6 8				1 1 2
E						E						E						E					7 1 2 6
TP6	1	2	3	4	5	TP8	1	2	3	4	5	TP9	1	2	3	4	5	TP10	1	2	3	4	5
A	1 1 1	3 5 3	7 1 1 5			A	1 3 1					A						A	2 4 2		4 1 0 4		
B		6 6 8				B	6 9 6		4 3 4			B	6 1 0 6	2 2 2	6 5 8			B	1 1 1	3 2 3	6 6 6		
C			4 2 6		2 2 2	C	6 1 4 4		3 1 4		2 1 2	C	5 6 4	1 1 1				C	7 4 8		5 8 5		7 6 7
D		5 3 8				D				5 5 6		D						D					
E						E						E		3 4 3		3 2 4		E					

Figure 10. Rankings of the segments on *NLop*, *BLoP* and *WLoP*  
 Note: The scores are shown in the segments above 75th percentiles on *NLop*. Those above 95th percentiles are emphasized with red borders.

Interestingly, the first and second modal segments on *NLop* (actually above the 95th percentiles) yielded fairly good agreement among the rankings on *NLop*, *BLoP* and *WLoP* as shown in Figure 10. Those with lesser modality (between 75th and 95th percentiles) produced good but weaker agreement. Also of interest is the lack of high modality of the rightmost segments (column 5) of Type B in contrast to Types A and C. Although not decisive, the difference might be attributable to the absence/presence of the right sub-area (see Figure 6).

## 4. DISCUSSION

To our knowledge, few previous studies examined fixation sequences with nested chunking. In view of the fact that the algorithms for fixation identification (see van der Lans et al., 2011) is still being developed, we should admit that our method for chunking is just a modest initial attempt. Its merit accrues from conceptual simplicity and clarity as well as ease of implementation.

We found that most chunks were singletons consisting of single fixations and that they were spread over all of the segments across TP excepting for one segment on one TP. It means that almost all segments attracted viewers attention for a given duration. However, more examination of the data is necessary to find how often the attention was drawn and how widely it was shared among viewers. Perhaps, a heatmap of singletons will help answer the first question, though not completely. For the second question, frequent pattern mining approach seems to be promising. Among others, we are currently testing PrefixSpan (Pei et al., 2001) to find out common patterns of singletons as well as chunks with multiple fixations. The mining task on the current fixation sequences will not be very complicated, since most frequent chunks were of length 1 or 2.

To the extent chunking reflects viewers continuous attention, the fixations within chunks would not widely spread over the display. A new chunk, on the other hand, would entail a jump to a new area, mostly akin to the preceding one provided the pre-saccadic attention during fixation (Goldberg et al., 1999). Indeed, we obtained the results compatible with these views. While the overwhelming majority of the within-chunk distance was zero or one, the between-chunk distance was relatively more dispersed.

A caveat is necessary about the segment-based distance as opposed to fixation-based one. Let us illustrate the point by Figure 11. Fixation,  $F_2$ , is close to  $F_3$  than to  $F_1$  as can be measured by the Euclidean distance. Yet, being in the segment the distance between  $F_2$  and  $F_1$  is zero, while that between  $F_2$  and  $F_3$  is one in terms of the vector length, i.e., segment-based distance employed in the present work.

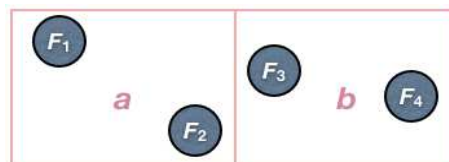


Figure 11. Fixations in Segments *a* and *b*

As stated in Introduction, the segment-based distance is useful in inferring sustained or recurrent interest (i.e., loops) in the same area-of-interest (AOI). If necessary, one can make close inspections of the selected by applying the Euclidian or other suitable distance.

We obtained some indications of layout effects all involving Type C: the lower participation rate of the segments in the homogenous chunks of length two (than Types A and B), the more frequent long between-chunk moves (than Types A and C) and the presence of the highly modal on the rightmost segments. To attribute the results to the location of the sub-areas of Type C (and A), we need studies with more rigorously controlled layout designs.

At times, we receive comments about the cultural influences of natural writings and readings. Three major styles are a) from left to right like in English, French and German, b) right to left in Arabic and Hebrew, c) co-practice of horizontal (left-to-right) and vertical in Japanese and Chinese (Taiwan). In light of the visual versatility of our subjects (all Japanese), we should refrain from drawing hasty interpretations. Even though web-pages in Japanese are written horizontally due to the constraints under HTML, people still show their preference for the vertical writings on iPad when it comes to novels and other contents.

So what is the best use of our approach including network analysis (Matsuda et al., 2009, 2010, 2011a, b, c), if nothing is conclusive yet. We believe the merits will be most appreciated by web-designers as well as other information-designers who are concerned about viewers reactions in relation to the designers' intent.

## REFERENCES

Book

Nielsen, J. & Pernice, K. 2010. *Eyetracking Web Usability*. Berkeley, CA: New Riders.

Journal

Goldberg, J. H. & Kotval, X. P. 1999. Computer interface evaluation using eye movements: methods and constructs. *International Journal of Industrial Ergonomics*, 24, 631-645.

Hayhoe, M. 2004. Advances in relating eye movements and cognition. *INFANCY*, 6(2), 267-274.

- Lorigo, L., Haridasan, M., Brynjarsdóttir, H., Xia, L., Joachims, T., Gay, G., Granka, L., Pellacini, F., and Pan, B. 2008. Eye tracking and online search: Lessons learned and challenges ahead. *Journal of the American Society for Information Science and Technology*, Vol. 59 (7), pp. 1041-1052.
- Martinez-Conde, S., Macknik, S.L., & Hubel, D.H. 2004. The role of fixational eye movements in visual perception. *Nature Reviews Neuroscience*, 5(3), 229-240. doi:10.1038/nrn1348.
- Matsuda, N., & Takeuchi, H. 2011a. Networks emerging from shifts of interest in eye-tracking records. *eMinds*, Vol.2(7), 3-16.
- Matsuda, N., & Takeuchi, H. 2011b. Joint analysis of static and dynamic importance in the eye-tracking records of web page readers. *Journal of Eye Movement Research*, Vol.4(1):5, 1-12.
- Matsuda, N., & Takeuchi, H. 2011c. Do heavy and light users differ in the Web-page viewing patterns?: Analysis of their eye-tracking records by heat maps and networks of transitions. *International Journal of Computer Information Systems and Industrial Management Applications*, (to appear in vol. 4).
- Melcher, D. 2008. Dynamic, object-based remapping of visual features in trans-saccadic perception. *Journal of Vision*, 8(14):2, 1-17.
- Parkes, M.B. 1992. *Pause and effect: An introduction to the history of punctuation in the west*. Hants, England: Scolar Press.
- Poole, A., & Ball, L.J. 2005. Eye Tracking in Human-Computer Interaction and Usability Research: Current Status and Future Prospects. In C. Ghaoui (Ed.) *Encyclopedia of Human-Computer Interaction*. Pennsylvania: Idea Group, Inc.
- Rayner, K. (1998). Eye Movements in Reading and Information Processing: 20 Years of Research. *Psychological Bulletin*, 124(3), 372-42.
- van der Lans, R., Wedel, M., & Pieters, P. 2011. Defining Eye-Fixation Sequences Across Individuals and Tasks: Binocular-Individual Threshold (BIT) Algorithm. *Behavior Research Methods*, 43(1), 239-257.
- Conference paper or contributed volume
- Matsuda, N., & Takeuchi, H. 2009. What will emerge from network analysis of shifts of interests of web page readers?: Another use of eye-tracking records. In *Proceedings of IADIS-IHC (Interface and Computer Interaction)*, pp.3-10; Algarve, Portugal, June, 20-22, 2009.
- Matsuda, N., & Takeuchi, H. 2010. Joint analysis of static and dynamic importance extracted from eye-tracking records of web page readers. In *Proceedings of IADIS2009 Interfaces and Human Computer Interaction (IHCI)*, paper #100, Freiburg, Germany.
- Pan, B., Helene Hembrooke, H.H., Gay, G., Granka, L.A., Feusner, M.K., Jill, K., & Newman, J.K. 2004. The determinants of web page viewing behavior: An eye-tracking study. In *Proceedings of Eye Tracking Research & Applications (ETRA2004)*, 147-154.
- Pei, J., Han, J.B., Mortazavi-Asl, B. & Pinto, H. 2001. "PrefixSpan: Mining sequential patterns efficiently by PrefixProjected pattern growth". In *Proceedings 2001 International Conference on Data Engineering (ICDE'01)*, pp. 215-224, Heidelberg, Germany, April 2001.
- Salvucci, D. D., & Goldberg, J. H. 2000. Identifying fixations and saccades in eye-tracking protocols. In *Proceedings of the Eye Tracking Research and Applications Symposium* (pp. 71-78). New York: ACM Press.
- Other papers
- Granka, L.A., Hembrooke, H.A. & Feusner, M.K. 2004. Corrlates of Visual Saliency and Disconnect: An Eye-tracking Evaluation. Unpublished research report, Cornell University Human-Computer Interaction Lab. <http://laura.granka.com/research.html>.
- Josephson, S. 2004. A Summary of eye-movement methodologies. FactOne: [http://www.factone.com/article\\_2.html](http://www.factone.com/article_2.html).
- Fulton, J.T. 2009. The saccades of the oculomotor system in vision. *Processes in Biological Vision* (Chap. 7) (online). <http://neuronresearch.net/vision/>.

# INTERACTIVE ART SYSTEM FOR MULTIPLE USERS BASED ON TRACKING HAND MOVEMENTS

Byunghyung Kim<sup>1</sup>, Danna Gurari<sup>1</sup>, Hough O'Donnell<sup>2</sup>, Margrit Betke<sup>1</sup>  
Boston University

<sup>1</sup> Image and Video Computing Group, Department of Computer Science,

<sup>2</sup> School of Visual Arts, College of Fine Arts

## ABSTRACT

We propose a camera-based human computer interface where hand movements directly influence the creation and manipulation of art in an interactive system. This unobtrusive, low-cost interface addresses limitations in previous interactive art systems by supporting multiple users simultaneously in an unconstrained, uncontrolled environment. We chose Markov Models for our computer vision algorithm to track multiple hands across multiple camera views in real time. We show how our algorithm can be applied as a tool for a team of users to interact with *Strange Attractors* and *Hidden Paintings*. *Strange Attractors* is an interactive art system that displays beautiful fractal structures whose location on a large screen reflects the users' hand movements. *Hidden Paintings* is an interactive game in which two almost identical versions of an artwork are displayed on a large screen and the movements of both hands of the users are superimposed as bounding boxes onto the display, allowing users to identify image regions in one version of the artwork that differ from the regions in the other. Experiments with twelve subjects show successful adoption of the proposed interactive art system and its two applications.

## KEYWORDS

Human computer interaction, camera-based human computer interfaces, computer vision, art, games

## 1. INTRODUCTION

Over the past decade, various systems have been explored that enable people to immerse themselves in images or videos and interact through their body movements. These systems have contributed to innovations in the performing arts, education, and entertainment [3][5][9][11][12][14][17]. This paper proposes a camera-based human computer interaction (HCI) system where movements of users' hands directly influence the creation and manipulation of art.

Related innovations in the performing arts include dance performances where a dancer's movements are tracked and used to steer a virtual representation in real time on a large monitor behind the dancer. In one setup, dancers' movements are detected by monitoring the location of wearable sensors [15]. This approach is limited because the sensors potentially impair the dancer's ability to move naturally. Another system uses 16 cameras to record body movement [12].

Related interactive art includes applications for improving the educational experience of children. One example is the *KidsRoom*, a media installation for guiding children through a story by interpreting their movements [3]. The installation was built as a bedroom-sized structure in a laboratory that contains a mattress and rug and is configured with the appropriate cameras, lighting, and screens. Another example is *SMALLab*, an immersive media environment used for kindergarten-through-12<sup>th</sup>-grade education to support active learning [2]. In this freestanding environment, teachers and groups of students interact with one another and their composed media worlds through free play, structured movement, and vocalization. A limitation for both systems is the complex setup required. *KidsRoom* requires students to visit the laboratory structure and *SMALLab* requires equipment for 3D-vision-based movement tracking, an audio microphone array, a video projector, and a multi-channel surround audio system.

The proposed HCI system is camera-based where a computer captures video of peoples' movements. The contribution of the proposed system, as an alternative to existing work, is a new, fun interface that addresses the limitations of the previous systems in a single system:

- **Multiple Users Supported** - several people can use the system simultaneously, encouraging interaction between users in addition to between the user and the system.
- **Unconstrained User Space** - any number of cameras can be used to increase the field of view and therefore allow for user movement across larger areas.
- **Unencumbered User Activity in Physical Space** – there are no wearable sensors; people can move naturally and interact with or create artworks using simple hand movements,
- **Easy to Use** – users can vary the direction and tempo of hand movements, making it easy and intuitive to create and interact with art pieces.
- **Simple installation** - minimal setup requirements makes the system suitable for everyday use (e.g., homes, offices, classrooms) and high-traffic spaces (e.g., galleries, museums).

Computer vision is used as the bridge between users' movements and art. The computer vision algorithm tracks hand motions (Section 2) and maps that movement graphically to the display in a way that is determined by the specific application (Section 3). The physical setup involves users, one or several low-cost web cameras, and a low-cost computer (Figure 1).

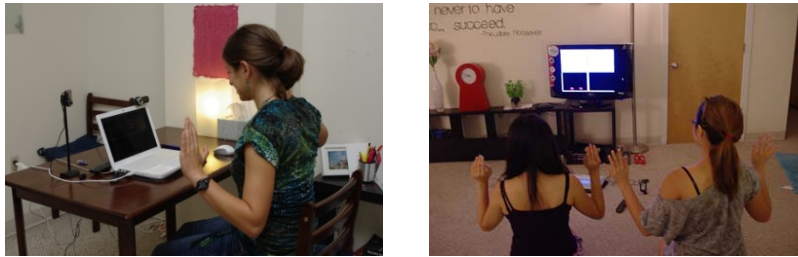


Figure 1. Interactive art system in use in two home environments with minimal hardware setup.

The key technical challenge in designing the proposed system was the development of a tracking algorithm that addresses the following requirements. The algorithm must have the ability to *simultaneously* track *multiple* hands of *different* users in *real time*. This means it must associate hands observed in any of the camera views with the hands tracked across previous views and discriminate against false hand detections. Many tracking methods, such as optical flow [1] and *CAMshift* [8], track objects based on features within shared fields of view between successive video frames making them unsuitable for tracking across multiple camera views. We employ a Markov Model to solve this data association problem, extending methods used at the University of Adelaide [6]. Our approach differs in that we track human hands instead of the human body and we permit human hands to move in various directions rather than just one direction. A Markov-Model-based approach is ideal because it uses observable states (not like hidden Markov models [7]), is computationally fast and accurate when tracking multiple objects, and distinguishes false hand detections. In addition, it handles tracking across multiple, non-overlapping camera views since discontinuities in visibility between two or more fields of view are accurately reflected in the Markov Model.

We experimented with two applications (Section 3). In *Strange Attractors*, our interface enables people to *create* beautiful fractal structures with quick energetic hand motions. A related application defines the form of the fractal structure with mouse input [10]. In *Hidden Paintings*, our interface enables players to actively *observe* artwork by pointing to different areas on pictures where expected objects are missing. Both applications encourage human-computer interactions based on a wide range of hand motions from quick, energetic gestures to carefully directed pointing. They demonstrate how the proposed interface can serve as a platform for engaging applications that support playful, creative interactions between users.

## 2. METHOD

A design goal of our interactive art system is that it can be used inexpensively for a wide range of applications in uncontrolled environments. It must therefore be resilient to real-world challenges such as changing illumination conditions, poor quality video, different backgrounds, many hand poses, a wide range

of skin tones, and occlusions. It must also provide real-time feedback. A flowchart of the proposed system is shown in Figure 2. The computer vision part of the system involves *detection of multiple hands* and *tracking of hands* and is followed by an *interactive art application*.

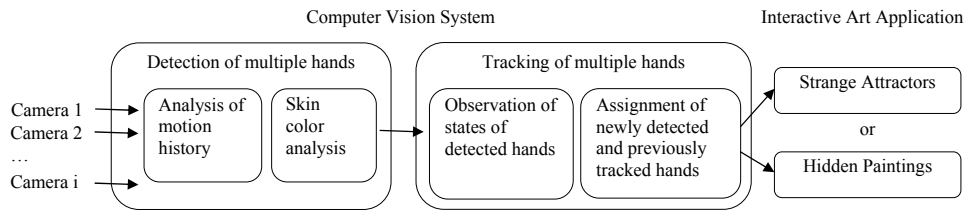


Figure 2. Overview of system for detection and tracking of multiple hands

## 2.1 Detection of Hands

Previous work demonstrated that motion and skin color are good indicators of hands in image sequences [18]. The proposed system therefore uses these cues for detecting hands. Motion can be analyzed in a computationally inexpensive way with a Motion History Image (MHI) [3]. MHI is a scalar-valued image where the pixel intensity is a function of how recently the motion occurred, with brighter pixels corresponding to the most current timestamps. The length of motion history retained is tunable. Our system captures a MHI from the past four images and creates a binary image with all static areas set to “0” (white in Fig. 3) and all movement areas set to “1” (blue in Fig. 3).

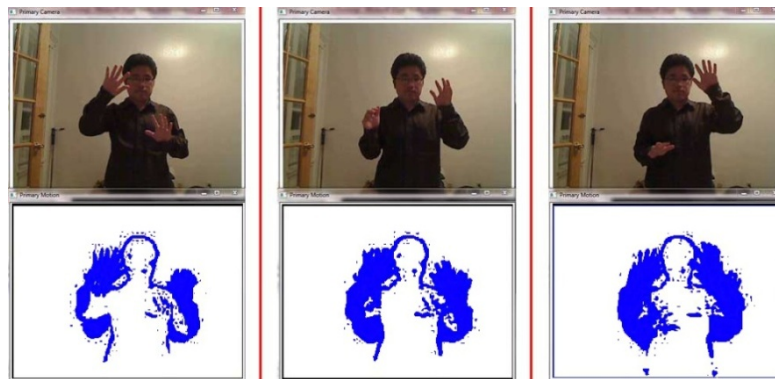


Figure 3. Visible images (top) and binary versions of corresponding motion history images (bottom)

In the binary image, clusters of pixels are grouped using connected component labeling and marked as hand candidates. Classification of hand candidates as hands is based on the dominant color and the size of the candidate. The skin color model  $F$  is defined during a training phase where a range of skin-colored hue values is determined. The hue component describes color in the form of an angle between 0 and 360 degrees ( $0^\circ = \text{red}$ ,  $120^\circ = \text{green}$ ,  $240^\circ = \text{blue}$ ,  $60^\circ = \text{yellow}$ ,  $300^\circ = \text{magenta}$ ). The skin-color model is defined as:

$$F(x) = \begin{cases} \text{True} & T_{\text{min}} \leq \text{hue value of pixel} \leq T_{\text{max}} \\ \text{False} & \text{otherwise} \end{cases}$$

where  $T_{\text{min}} = 70^\circ$  and  $T_{\text{max}} = 256^\circ$  are the minimum and maximum hue values, respectively, that indicate skin color. For each hand candidate region, if the peak value of its hue histogram falls within the pre-defined range of skin colored hues and its size is greater than a pre-defined minimum size, it is classified as a hand. This feature-based approach is ideal for our application because it is simple, fast, and accurate.

## 2.2 Multiple-Hand Tracking across Several Non-overlapping Camera Views

The next challenge addressed by the system is tracking the left and right hands of the interacting users that are in the field of view of one of the cameras. The proposed solution is based on a Markov Model which characterizes how each hand will move in the future using knowledge of how the hand has moved in the past. At time  $t$ , our system uses  $i$  cameras with non-overlapping views  $C_1, \dots, C_i$ , has detected  $j$  hands  $D_1, \dots, D_j$  in



these  $i$  views, and maintains a list of  $m$  tracked hands  $H_1, \dots, H_m$ . The task to assign the  $j$  currently detected hands to the  $m$  previously tracked hands is solved at each time step  $t$  by selecting the most probable set of assignments based on the transition probabilities of the Markov model. In the proposed system, an observable Markov Model state is a region in the camera view. Each video frame is divided into 16 states resulting in 32 states for a two camera setup as shown in Figure 4. The number of states must be selected such that a balance is struck between tracking accuracy and system speed. More states return improved tracking accuracy at the cost of slower performance and vice versa. The parameters defining the Markov Model are the state transition probability matrix and initial state probability vector, which are estimated during a training phase. When the proposed system is in the tracking phase, the transition probabilities of the Markov Model guide the tracking process.



Figure 4. Each image is divided into 16 states resulting in 32 states across two camera views

**Training Phase:** Estimates for the initial probability vector are assigned such that each state has an equal probability. The Markov Model used is ergodic, which means that any state can be reached from any other state in a finite number of steps. The state transition probability matrix is therefore initialized with equal probabilities. During training, its entries are adjusted using a series of observations collected as users move their hands in front of the cameras. The resulting matrix stores the probabilities that hands will move from a given region in an image (state) to another region in this image or in the image recorded by an adjacent camera (state). The transition probabilities trained for state 5, which represents a region in the center of the first image, are shown in Figure 5. The data shows that the hands did not move far during training and most likely appeared in surrounding image regions. This reflects the nature of human movement. As a result, in practice, we only need to be concerned with a localized area for state transitions.

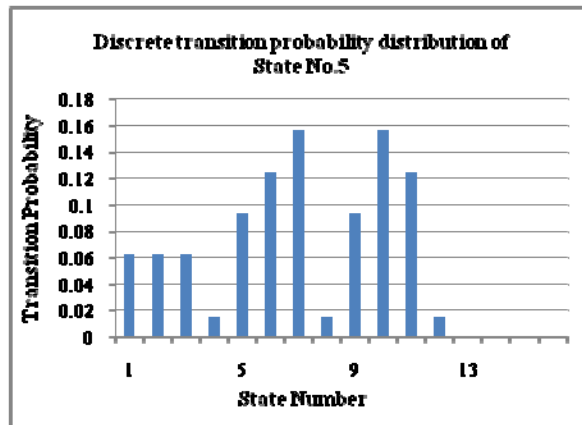


Figure 5. Discrete transition probability distribution of a trained Markov Model for state 5

**Tracking Phase:** During tracking, the system assigns the  $j$  detected hands to the  $m$  previously tracked hands. At each time step  $t$ , every detected, unmatched hand is allocated to its corresponding state using the location of its center in the image where it appears for the mapping. For each detected hand  $D$ , the most probable assignment to a tracked hand is calculated. In this calculation, a sliding time window of duration  $k$  is used. For each tracked hand  $H$ , the likelihood that it ended up in the image region (state) where  $D$  has been detected is computed by multiplying the probabilities of the last  $k$  transitions of hand  $H$ , ending with the transition into  $D$ . The assignment process has quadratic time complexity  $O(j m)$  at each time step.

### 2.3 *Strange Attractors: Creating Art by Gesturing with Both Hands*

The proposed application, *Strange Attractors*, is an interactive art system that displays fractal structures whose location, shape, and color reflect the users' hand motions (Figure 6). It was designed for an installation in the lobby of a high-traffic dormitory building at our university. Hand movements captured by the primary camera are passed as parameters driving the display of a new fractal structure. The fractal structure is generated by a dynamic system described by a differential equation [10]. Hand movements captured by the secondary camera change the dominant color of the rendered fractal structures. The *Strange Attractors* application encourages users to playfully interact with the computer and learn how they can manipulate the apparent chaos of displayed fractal structures.

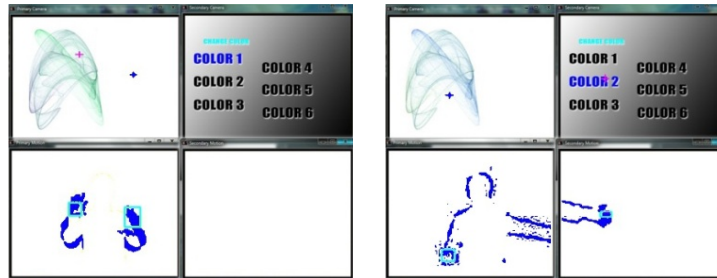


Figure 6. *Strange attractors*: The interface contains windows that displays the art (top left), the controller of the attractor color (top right) and the tracking results (cyan boxes) overlaid on the motion history image (blue) for the two camera views (bottom). The positions of the tracked hands in the primary camera view (bottom left) are input parameters to create the shape and location of the attractor. Tracked hands in the secondary camera view (bottom right) change the attractor color (here from green to blue).

### 2.4 *Hidden Paintings: Interacting with Art by Gesturing with Both Hands*

A second application that uses our proposed multiple-hand tracking system is an interactive game called *Hidden Paintings*. It is an educational, media-based application for locating differences in two almost identical versions of a picture using hand gestures. Multiple players can play together to find the hidden objects in the picture. A screen shot of a game is shown in Figure 7.



Figure 7. *Hidden Paintings*: The interface contains windows that display two versions of a picture\* and tracking results. The detected hand locations (cyan) are shown overlaid onto the binary version of the motion energy image. The tracked locations of the left and right hands are mapped to corresponding positions in the original version of the picture (blue and green rectangle). The goal of the user is to move the rectangles onto the objects that do not appear in the other version of the picture, here the fence and the red apple. (\* The original picture was obtained from Google Image Search and manipulated with image editing software to create a version with “hidden” objects.)

## 3. EXPERIMENTS AND RESULTS

Two sets of experiments were conducted using two web cameras streaming visible video of two adjacent, non-overlapping views. Cameras were positioned approximately 25-40 cm apart from each other with their optical axes aligned approximately parallel to one other. In the first set of experiments, the computer vision system was tested; in the second set of experiments, the focus of the tests was the HCI component.

### 3.1 Evaluation of Multiple-User Hand Tracking across Camera Views

For these initial experiments, learning of the Markov Model was performed off-line through manual annotations of three video sequences in which users were asked to move "just for fun." As a result, the trained Markov model represents arbitrary movements in space. In one system set-up, the hands of a single participant moved freely within the field of views of two cameras. Experiments determined that the system successfully tracked motion of hands between the two fields of views. In a second system setup, the system successfully tracked two people's four hands which moved freely within the fields of view of two cameras (Figure 8). The main challenge of this setup is assigning hands correctly when they are close to each other or occluding each other (Figure 8, row 6). This problem can be solved by dividing each field of view into more regions resulting in additional states in the Markov Model. The downside of increasing the number of states is slower system performance. In testing, by increasing the number of total states from 50 to 98, the overall tracking accuracy increased from 68.95% to 86.64% based on 1430 observations.



Figure 8. The system tracks the hands of two participants who move them in either camera's field of view.

### 3.2 Evaluation of the Interactive Art System

Experiments were conducted with twelve participants playing *Strange Attractors* and *Hidden Paintings*, both which are based on camera-detected hand movements. Participants included Caucasians, Asians, and African Americans with different skin color tones. Their ages ranged from 5 to 60. Participants cooperated with the system by not creating occlusions of hands. The experiments were run in a variety of settings including public places such as coffee shops and private places such as apartments.

One goal was to test whether the system sufficiently would track quick, energetic motions as well as precisely articulated motions as users move naturally in a variety of environments. Participant feedback about *Strange Attractors* was that the fractal structures looked beautiful, but that it was difficult to recognize the connection between the shape of the structures and the users' motions. Fittingly, users made a range of lively hand movements as they tried to learn how the system would respond to their actions. *Hidden Paintings*, in contrast, encouraged defined hand movements from users. The experiments showed users enjoyed creating fractal art with *Strange Attractors* and playing the *Hidden Paintings* game, indicating system performance was subjectively sufficient for users as they moved in a variety of ways. Moreover, consumer-level cameras provided adequate video quality and our system successfully detected various skin tones.

In a quantitative set of experiments, we compared the time it took participants to complete a task with camera-detected hand movements with the time it took them to perform the same task with the traditional computer mouse. The assigned task was to finish eight stages of the *Hidden Paintings* game. The total time per stage was calculated for each player and the results were grouped according to the players' ages (Figure 9). Ages for each group are 5-19 for Group 1, 20-40 for Group 2, and 41-60 for Group 3. A common factor for most players was that less time was taken with experience; the interaction time typically decreased from game 1 to 5. It is concluded that people become accustomed to the interface through repeated experiences.

Another interesting result is that younger generations comparatively performed the tasks faster than older generations, independent of the interface. We also measured that the time to finish the game was lower for mouse-based interaction than for camera-based interaction by approximately a factor of 2/3. In an extreme case, it took a participant more than twice as long to finish a game with the camera-based interface than with the mouse-based interface.

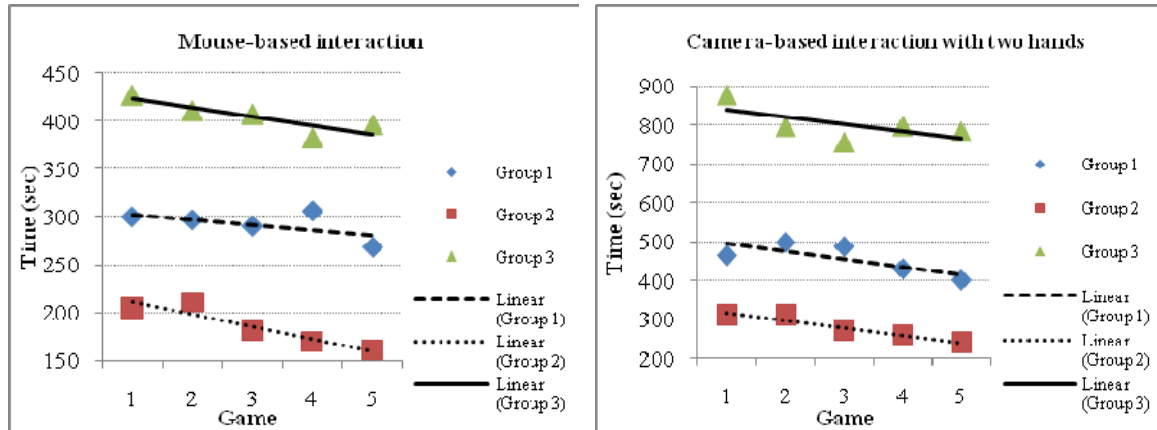


Figure 9. Timing results for mouse-based interactions (left) and camera-based interaction with two hands (right).

We also computed the ratio of average mouse-based and camera-based task-completion times per game and per group in two experiments. The plot in Figure 10 indicates that the rate of the task-completion time stayed flat for groups 2 and 3 and increased for group 1. This means that the younger participants were able to speed up their interaction with the previously unknown camera-based interface much more than with the common mouse-based interface. The older participants might have also experienced this speed-up if they had more practice. The question remains open whether training could result in camera-based interaction times that are similar to mouse-based interactions. The difference in interaction times becomes irrelevant for art installations where user access to a computer mouse is impractical (e.g., shopping mall installations).

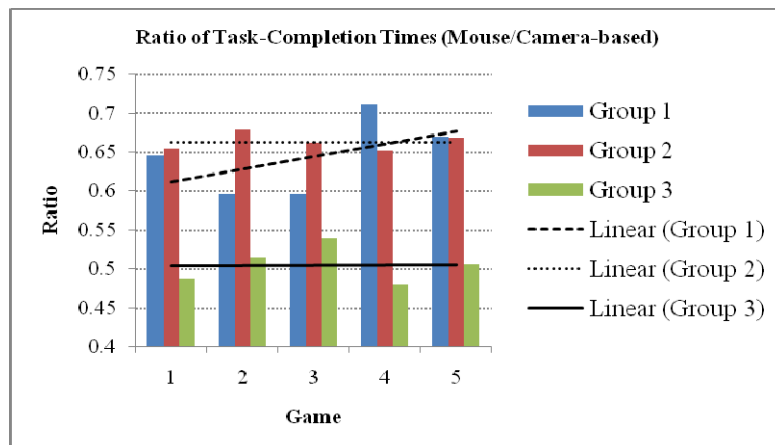


Figure 10. Comparison of mouse-based and camera-based task-completion times for *Hidden Paintings*

#### 4. CONCLUSIONS AND FUTURE WORK

The proposed interface enables users to create and manipulate artwork with movements of their hands. It sufficiently tracks and interprets a wide range of hand motions from quick, energetic gestures to carefully directed pointing. Users may interact with each other while moving naturally in a variety of environments. The interactive system is low cost and easy to use, requiring only inexpensive web cameras and computers

with minimal setup. It currently supports two example applications, *Strange Attractors* and *Hidden Paintings*. Participants of HCI experiments with the interface enjoyed creation and manipulation of artworks.

Future work will be focused on three tasks. First, the computer-vision part of the system will be enhanced so that the creation of the Markov Model is automated. The system may then be used more easily in new environments. Second, we plan to add other applications to the system that enable users to create and manipulate artworks. Third, we will study user reaction to the system when it becomes an art installation in the lobby of our university dormitory.

## ACKNOWLEDGEMENTS

The authors thank the test subjects for their time and gratefully acknowledge NSF funding (IIS-0910908, IIS-0855065, and IIS-0713229).

## REFERENCES

- [1] Adiv, G., 1985. Determining Three-Dimensional Motion and Structure from Optical Flow Generated by Several Moving Objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*: 7(4):384-401.
- [2] Bircheld, D. et al., 2006. Sound and interaction in K-12 mediated education. *Proceedings of International Computer Music Conference*. New Orleans, USA.
- [3] Bobick, A. F. et al., 1999. The KidsRoom: A Perceptually-Based Interactive and Immersive Story Environment. In *PRESENCE: Teleoperators and Virtual Environments*, Vol. 8, No. 4, pp. 369-393.
- [4] Bobick, A. F. and Davis, J. W., 2001. The Recognition of Human Movement Using Temporal Templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*: 23(3):257 -267.
- [5] Castellano, G. et al., 2007. Expressive Control of Music and Visual Media by Full-Body Movement. *Proceedings of the Conference on New Interfaces for Musical Expression*. New York, USA, pp. 390-391.
- [6] Dick, A. and Brooks, M., 2004. A Stochastic Approach to Tracking Objects Across Multiple Cameras. *Proceedings of Australian Conference on Artificial Intelligence*. Sydney, Australia, pp. 160-170.
- [7] Elmezain, M. et al., 2008. A Hidden Markov Model-based continuous gesture recognition system for hand motion trajectory. *Proceedings of 19th International Conference on Pattern Recognition*. Florida, USA, pp. 1-4.
- [8] Exner, D. et al., 2010. Fast and robust CAMShift tracking. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition Workshops*. San Francisco, USA, pp. 9-16.
- [9] Freeman, W. T. et al., 1998. Computer vision for interactive computer graphics. *IEEE Computer Graphics and Application*, 18(3): 42-53.
- [10] Grant, S., 2008. Attractors and Fractals. <http://www.cc.gatech.edu/~phlosoft/attractors>.
- [11] Hamalainen, P. And Hoysniemi, J., 2002. A Computer Vision and Hearing Based User Interface for a Computer Game for Children. *Proceedings of the 7th ERCIM Workshop "User Interfaces For All"*. Paris, France, pp. 23-25.
- [12] James, J. et al., 2006. Movement-based interactive dance performance. *Proceedings of the 14th International Conference on Multimedia*. Santa Barbara, USA, pp. 470-480.
- [13] Kakumanu, P. et al., 2007. A Survey of Skin-Color Modeling and Detection Methods. *Pattern Recognition*, 40(3): 1106-1122.
- [14] Levin, G., 2006. Computer Vision for Artists and Designers: Pedagogic Tools and Techniques for Novice Programmers. In *Journal of Artificial Intelligence and Society*, Vol. 20, No. 4, pp.462-482.
- [15] Park, C. et al., 2006. A wearable wireless sensor platform for interactive art performance. *Proceedings of IEEE International Conference on Pervasive Computing and Communications*. Pisa, Italy.
- [16] Pickover, C., 1995. *The Pattern Book: Fractals, Art, and Nature*. World Scientific Publishing Company.
- [17] Sparacino, F. et al., 2000. Media in performance: Interactive spaces for dance, theater, circus, and museum exhibits, *IBM Systems Journal*, 39: 479-510.
- [18] Zhu, Q. et al., 2004. Adaptive learning of an accurate skin-color model. *Proceedings of the 6th IEEE International Conference on Automatic Face and Gesture Recognition*. Seoul, Korea, pp. 37-42.

# FEATURE EXTRACTION BASED ON EMPIRICAL MODE DECOMPOSITION AND BAND POWER APPROACHES FOR MOTOR IMAGERY TASKS CLASSIFICATION

Dalila Trad<sup>1</sup>, Tarik Al-ani<sup>2</sup>, Eric Monacelli, Stéphane Delaplace and Mohamed Jemni<sup>1</sup>

UVSQ, LISV, 10/12 Avenue de l'Europe, 78140 Vélizy, France

<sup>1</sup>University of Tunis, UTIC, 5, avenue Taha Hussein, B.P. 56 Bab Menara 1008 - Tunis Tunisia

<sup>2</sup>ESIEE-Paris, Dep. Informatique, Cité Descartes, 93162 Noisy-Le-Grand, France

## ABSTRACT

Brain-Computer interface (BCI) is a system that allows paralyzed patients to communicate with their environment. In this work we investigate a nonlinear and non stationary framework for feature extraction and classification of *Electroencephalogram* (EEG) signals in order to classify motor imagery tasks in BCI. The feature extraction approach is based on the *empirical mode decomposition* (EMD) and *band power* (BP). The EMD generates several sets of stationary component series, *Intrinsic Mode Functions* (IMFs). These IMFs are then analyzed with the *power spectral density* (PSD) to study the active frequency ranges corresponding to the motor imagery tasks of each subject ( $\square$  and  $\square$  rhythms) in the EEG signals. Finally, the data were reconstructed with only the specific IMFs and then the band power (BP) is employed on the new database. Once the new feature extraction is applied, the classification of motor imagery is done using *Hidden Markov Models* (HMMs). The obtained results showed that the EMD method allows the extraction of the most reliable features from EEG and that the obtained classification rate is higher and better than using only the direct BP approach. Then, the recognized mental tasks may be translated into a low-level command in order to help persons with myopathic diseases or *muscular dystrophy* (MD) to move a joystick to a desired direction.

## KEYWORDS

Brain-Machine Interface (BMI); Brain-computer interface (BCI); Electroencephalogram (EEG); Signal processing; Classification; Assistive technology.

## 1. INTRODUCTION

*Brain-Computer interface* (BCI) is a direct communication pathway between a brain and an external device. The major goal of the BCI research is to develop systems which help disabled users to communicate with other people in order to control artificial limbs or their environment (Wolpaw et al., 2002), to control an electric device like moving a cursor (Wolpaw et al., 1994) or to control hand prosthesis (Guger et al., 1994) by the imagination of left and right hand movement.

A BCI system is represented as a system in a continuous closed loop generally composed of six steps, Figure 1: Brain activity measurement, pre-processing, feature Extraction, classification, translation into a command and feedback (Mason and Birch, 2003).

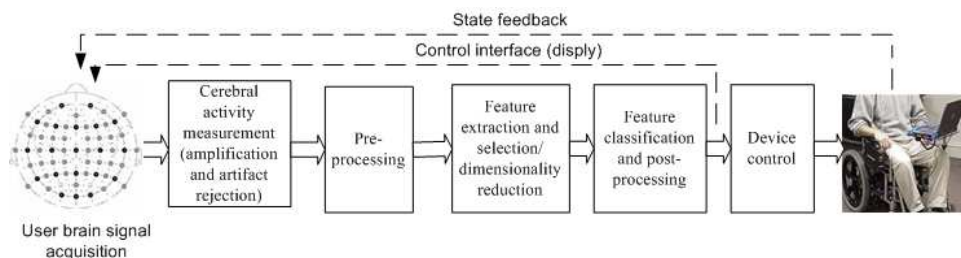


Figure 1. General functional model of a BCI System.

One major challenge of *BCI* is thus to extract reliable information (features) from noisy *EEG* data, i.e. step 3. These features can then be used in step 4 in order to classify the user's mental state. Physiological studies (Pfurtscheller, 1999; Pfurtscheller and Neuper, 2001) showed that the rhythms  $\mu$  (8-12 Hz) and  $\beta$  (13-30 Hz) are the main relevant information for discriminating motor activity. A common approach in *BCI* is thus to extract the *band powers* of the rhythms  $\mu$  and  $\beta$  from the *EEG* signal and use them as a classification features. Several common *band power* techniques were employed in the *BCI* literature. Herman et al. (Herman et al., 2008) demonstrated that the Yule and Welch *PSD* approaches, mainly dominate the other studied approaches. These approaches are based essentially on some linearity and stationarity hypothesis such as the use of *fast Fourier transform (FFT)* spectrum in short time of a segment of data. A method like FFTs is based on a linearity and piecewise stationarity hypothesis (Penzel and Conradt, 2000). However, the *EEG* is naturally nonstationary and nonlinear (Lo et al., 2009). Thus, it is more convenient to apply a nonstationary and nonlinear signal processing approach on the *EEG*. Recently, Huang et al. (Huang et al., 1998) proposed a technique called the *empirical mode decomposition (EMD)* for nonlinear and nonstationary time series data. The *EMD* is a data driven approach that can be used to decompose adaptively a signal into a finite number of mono-component signals, which are known as *intrinsic mode functions (IMFs)* or *modes*. It considers signals at their local oscillations, but they are not necessarily considered in the sense of Fourier harmonics. Their extraction is non-linear, but their recombination for exact reconstruction of the signal is linear. In this work, we apply first the *EMD* on the *EEG* signals and then we apply the Welch-based *band power* for feature extraction in order to extract the reliable information of *EEG* corresponding to some motor imagery tasks. Based on these features, the classification of the mental tasks was done using a nonlinear classifier known as *hidden Markov models (HMMs)* (Rabiner, 1989).

The aims of this work is to use a control architecture based on a fusion of a non-invasive *BCI* system with a *haptic* device adapted to assist a person with myopathic diseases who has lost partially the use of his hands. It is a reinforcement strategy of the detection mechanism of intent based on the processing of *EEG signals*. This strategy offers two advantages: a. It retains the use of conventional *human-machine interface* devices for people with severe disabilities (*myopathy*, *Tetraplegia* or *Cerebral Palsy*). b. It couples these interfaces with a minimal *BCI* system as this system is dedicated only to information related to the hands. Thus, we implemented in this work a Simulink/MathWorks model that translates, on-line, the different mental tasks *EEG* signals into low-level commands to reinforce the desired movement of a joystick.

## 2. DATA BASES & METHODS

### A. EEG Data

Two motor imagery *EEG* data corresponding to three subjects were used in this work.

For subject 1, the *EEG* data corresponds to an experiment with four sessions "run1234" acquired by Guger Technologies (<http://www.gtec.at/>). Each session contains 40 trials: 20 trials for right hand movement imagination and 20 trials for left hand movement imagination, where each trial contains 2048 samples. The timing of this experiment is shown in Figure 2.

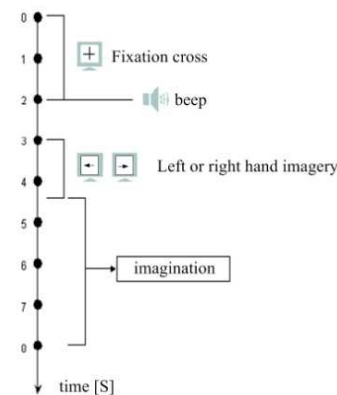


Figure 2. Paradigm: Timing of one trial in the experiment.

The mechanism is described as follows: after two seconds a warning stimulus was given of a ‘beep’. From second 3 until second 4.25 an arrow pointing left or right was shown on the monitor after that the subjects were instructed to imagine a left or right hand movement depending on the direction of the arrow. For subjects 2 and 3, the *EEG* data was recorded in our Department using the biosignal amplifier “g.USBamp-Gtec” (<http://www.gtec.at>). Subject 2 (abled female aged 24 years) and subject 3 (abled male aged 30 years) were instructed to imagine a left or right hand movement depending on the direction of the arrow. The same preceding experiment and data structure were obtained.

The experiment data were sampled at 256 Hz and filtered in the range of 0.5 and 30Hz. A notch filter was used to suppress the 50Hz power line interference. Two bipolar recordings overlying the left and right sensorimotor area were obtained by two electrodes C3 and C4 placed according to the international 10/20 system (Niedermeyer and Lopes da Silva, 2004).

### B. Feature extraction

In brain-computer interface, feature extraction has an important role. In our work, we proposed a feature extraction method based on *EMD* and *BP* in order to extract the relevant features of the *EEG* signal for left or right motor imageries recognition.

1) Empirical mode decomposition (*EMD*) approach: The traditional *EMD* was recently proposed (Huang et al., 1998) as an adaptive time-frequency data analysis method. It is defined by an algorithm based on an empirical framework. The basic *EMD* is defined by a process called *sifting* to break down any multimodal signal to a sum of basis components called *intrinsic mode functions* (*IMFs*). The *IMFs* are zero-mean AM-FM signals which must satisfy two conditions: the first one is that the number of extrema and that of zero-crossing must differ at most by one; the second one is that the mean value between the upper and lower envelopes are equal to zero at any point. Conceptually, the establishment of this method is quite simple: one needs to consider a signal at its local oscillation level, remove the fastest oscillation and iterate the process on the residue considered as a new signal. At the end of the sifting processes, a given signal  $x(t)$  can be written as a sum of a finite number of *IMFs*,  $I_m(t)$ ,  $m = 1, 2, \dots, M$ , and a final residue  $r_M(t)$ :

$$x(t) = \sum_{m=1}^M I_m(t) + r_M(t).$$

The decomposition is stopped at step  $M$ , if either the residue  $r_M(t)$  is a mono-component signal or has only 2 extrema (Huang et al., 1998). The *stopping criterion* must be set to ensure that the obtained signal satisfies the properties of an *IMF* while limiting the number of iterations. For more details about the different steps of the *sifting* process for the calculation of the *IMFi* as well as the *stopping criterion* definition see (Huang et al., 1998). Since the decomposition into *IMFs* is based on the local characteristic time scale of the data, it applies to nonlinear and non-stationary processes. The *IMFs* admit well-behaved *Hilbert transforms* (*HT*) (Gabor, 1946) and they satisfy the following properties: they are symmetric, different *IMFs* yield different instantaneous local frequencies as functions of time that give sharp identifications of embedded structures. The decomposition is done linearly or non-linearly depending on the data. This complete and almost orthogonal decomposition is empirically realized by identifying the physical local characteristic time scales intrinsic to these data, which is the lapse between successive extrema.

2) Band powers (*BP*): The features may be extracted from the *EEG* signals by estimating the power distribution of the *EEG* in predefined frequency bands. In general, the band power is estimated by digitally band-pass filtering the data, squaring and averaging over consecutive samples according to a given window size. Pfurtscheller et al. (Pfurtscheller et al., 1997) used the *BP* and demonstrated that for each subject, different frequency components in the  $\mu$  (9 - 14 Hz) and  $\beta$  (18 - 26Hz) bands were found which provided best discrimination between left and right hand movement imagination.

In this work, we propose a direct nonlinear approach to extract the more relevant *IMFs* corresponding to the different frequency components in the  $\mu$  and  $\beta$  bands and then obtain the Welch-based *BP* and use them as features for mental task classification. We applied the *EMD* method on the *EEG* data defined in section II-A. The *EEG* data for each subject are composed of 80 trials corresponding to left hand movement imaginations (C4) and 80 trials corresponding to right hand movement imaginations (C3). Figure 3 (a) and Figure 4 (a) show the result of one-trial (left hand movement imagination) *EMD* decomposition for subject 2 in the channels C3 and C4 respectively. Each channel is decomposed into 10 *IMFs* and one residue. Welch’s method was applied to analyze the different characteristics of each *IMF*, (Welch, 1967). This method estimates the *PSD*, it was applied to each *IMF* to calculate and find the active frequency bands such as the  $\mu$



and  $\beta$  rhythms. Figure 3 (b) and Figure 4 (b) show the *PSD* in each *IMF* in the channels C4 and C3 respectively.

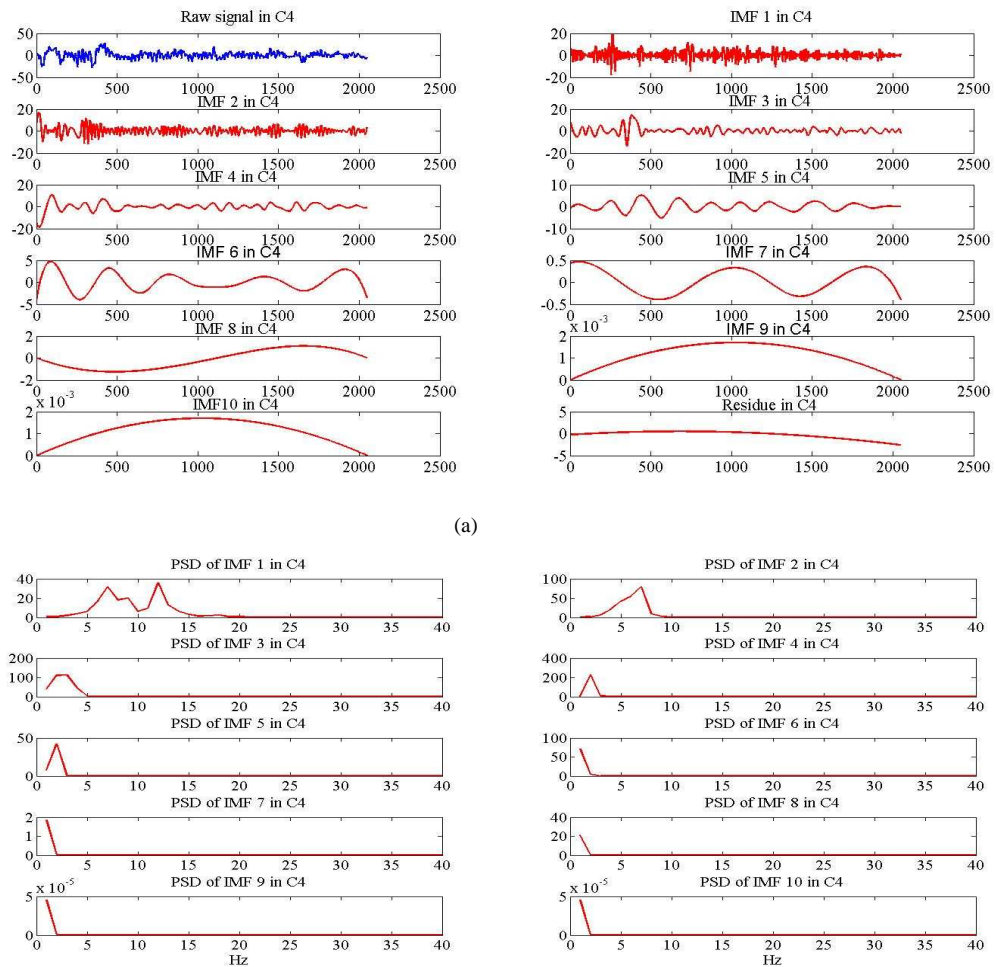


Figure 3. *IMFs* and *PSD* in C4 during left hand movement imagination.. (a) The raw signal of subject 2, the ten *IMFs* and the residue as a function of sample index. (b) *PSD* (dB/Hz) of each *IMF* in (a).

We can notice that the characteristics of the active frequency bands corresponding to  $\mu$  and  $\beta$  are located only in *IMF1*, *IMF2* on C3 and C4. Concerning subject 1, the active frequency bands are located only in *IMF1*, *IMF2* and *IMF3* on C3 and C4. Therefore, the new signal is reconstructed by keeping only the two first *IMFs* for subject 2 and only three first *IMFs* for subject 1. Then, band power was applied for the new signal.

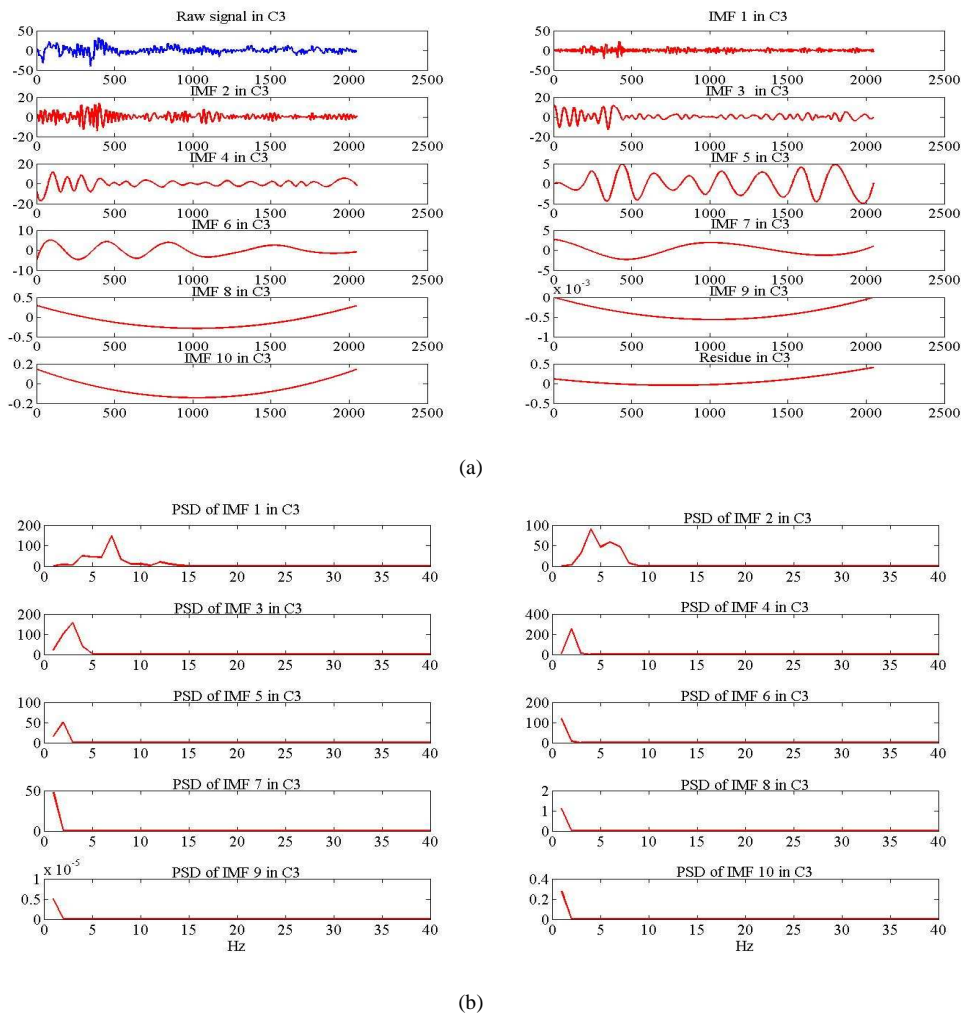


Figure 4. *IMFs* and *PSD* in C3 during left hand movement imagination. (a) The raw signal of subject 2, the ten *IMFs* and the residue as a function of sample index. (b) *PSD* (dB/Hz) of each *IMF* in (a).

### C. Classification

The *HMM* method is very efficient nonlinear technique used for the classification of time series (Rabiner, 1989). It necessitates two stages: a training stage where the stochastic process models are estimated through extensive observation corpus and decoding or detection stage where the model may be used off/on-line to obtain the likelihoods of the given test sequence evaluated by each model. A *HMM* is defined by the following compact notation to indicate the complete parameter set of the model  $\lambda = (\mathbf{P}, \mathbf{A}, \mathbf{B})$ , where  $\mathbf{P}$ ,  $\mathbf{A}$  and  $\mathbf{B}$  are the initial state distribution vector, matrix of state transition probabilities and the set of the observation probability distribution in each state, respectively. This set of parameters is defined by  $\Pi = [\pi_1, \pi_2, \dots, \pi_N]$ ,  $\pi_i = P(q_1 = s_i)$ ,  $\mathbf{A} = [a_{ij}]$ ,  $a_{ij} = P(q_{t+1} = s_j | q_t = s_i)$ , where  $1 \leq i, j \leq N$ ,  $s_i, s_j \in S$ ,  $S = \{s_1, s_2, \dots, s_N\}$ ,  $t \in \{1, 2, \dots, T\}$ . The observation at time (or index)  $t$ ,  $\mathbf{O}_t$ , is considered in this paper as continuous  $\mathbf{O}_t \in \mathbb{R}^K$ . For a continuous observation, the state conditional probability of the observation  $b_i(\mathbf{O}_t)$  may be defined by a finite mixture of any log-concave or elliptically symmetric probability density function (pdf), e.g. Gaussian pdf, with state conditional observation mean vector  $\boldsymbol{\mu}_i$  and state conditional observation covariance matrix  $\boldsymbol{\Sigma}_i$ . In this paper we consider only a single Gaussian pdf, so  $\mathbf{B}$  may be defined as  $\mathbf{B} = \{\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\}$ ,  $i = 1, 2, \dots, N$ . At each instant of time  $t$ , the model is in one of the states  $i$ ,  $1 \leq i \leq N$ . It outputs  $\mathbf{O}_t$  according to a density function  $b_i(\mathbf{O}_t)$  and then jumps to state  $j$ ,  $1 \leq j \leq N$  with probability  $a_{ij}$ . The state transition matrix defines the structure of the *HMM*. The model  $\lambda$  may be obtained off-line by a given training procedure. In practice, given the observation sequence  $\mathbf{O} = \{\mathbf{O}_1, \mathbf{O}_2, \dots, \mathbf{O}_T\}$ , and a model  $\lambda$ , the *HMMs* need three fundamental problems to

be solved (Rabiner, 1989). The first problem is how to calculate the likelihood  $P(O|\lambda)$ ? The solution to this problem provides a score of how  $O$  belongs to  $\lambda$ . The second problem is how to determine the most likely state sequence that corresponds to  $O$ ? The solution to this problem provides the sequence of the hidden states corresponding to the given observation sequence  $O$ . The third problem is how to adjust the model  $\lambda$  in order to maximize  $P(O|\lambda)$ ? This is the problem of estimating the model parameters given a corpus of training observations sequences. Problems 1 and 2 are solved using the forward and the Viterbi algorithms respectively. However, in our work, the probability along the Viterbi path is used to estimate  $P(O|\lambda)$ . Problem 3 is solved during the training phase using either a conventional algorithm such as the Baum-Welch algorithm (Rabiner, 1989). Our training scheme is based on Baum-Welch algorithm and the *Bayesian Inference Criterion (BIC)* (Schwarz, 1978; Helmy et al., 2008) using multiple feature sequences (trials). This scheme makes the training procedure independent of the initialization problem and the a priori knowledge of the number of states in each *HMM* needed in the Baum-Welch training algorithm.

*D. Results*

Table 1 shows the *HMMs*-based classification results using the two feature extraction methods direct *BP* and *EMD + BP*. For each subject, the *EEG* data contains 160 trials and each trial lasts 8 seconds as shown in Figure 2 (a set of 80 trials for left hand movement imaginations and a set of 80 trials for right hand movement imaginations). Each set of movement imagination data was divided into two subsets for each mental task movement (40 trials for *HMMs* training and 40 trials for test). For the *HMMs* training and test steps, each trial (C3 or C4 feature sequence) is composed of  $T = 2048$  samples, where each sample is a scalar value ( $K=1$ ). In these subsets, we considered only the imagination period: 4 seconds to 8 seconds (see Figure 2). We trained one *HMM* for each subject and for each mental task (right hand or left hand movement) using the data corresponding to that subject and that mental task. The number of states in each *HMM* was determined automatically by the BIC. This number is 2 in each *HMM*. Table 1 shows an improvement in the identification of "left" imagination by all subjects is obtained using *EMD + BP*. However, Table 1 shows that the identification of "right" imagination by subjects 1 and subject 2 was identical for the two feature extraction methods. The improvement of the identification of the two "right" and "left" imaginations for subject 3 shows that the result for subjects 1 and 2 could be an odd coincidence.

Table 1. Confusion matrices for left and right hand movements of subjects 1, 2 and 3 using the two feature extraction methods (BP and EMD+BP) and *HMMs*.

Subjects	Feature Extraction Method	Actual task	Predicted right	Predicted left
Subject 1	<i>BP</i>	Right	85.0%	15.0%
		Left	10.0%	90.0%
	<i>BP + EMD</i>	Right	85.0%	15.0%
		Left	2.5%	97.5%
Subject 2	<i>BP</i>	Right	72.5%	27.5%
		Left	47.5%	52.5%
	<i>BP + EMD</i>	Right	72.5%	27.5%
		Left	20.0%	80.0%
Subject 3	<i>BP</i>	Right	65.0%	35.0%
		Left	22.5%	77.5%
	<i>BP + EMD</i>	Right	80.0%	20.0%
		Left	22.5%	57.5%

Table 2 shows the means of the recognition percentage rates given in Table 1 for the three subjects. It can be seen that the mean classification rates in the two types of movement were increased by 5% with the *EMD + BP* with respect to the direct *BP* method.

Table 2. Means of the recognition percentage rates for the three subjects given in Table 1.

Mouvement	Right (means for Subjects 1, 2 and 3)		Left (means for Subjects 1, 2 and 3)	
	<i>BP</i>	<i>EMD + BP</i>	<i>BP</i>	<i>EMD + BP</i>
HMM classifier	74.167 %	79.167 %	73.33 %	78.33 %

*E. Translation into a command*

Once the motor imagery is identified, a command may be associated to this mental task in order to control a machine such as a hand prosthesis (opening/closing) (Guger et al., 1994) or a robot (Kbler et al., 2006). In this work, we constructed a new Simulink/MathWork model to translate on-line the *EEG* signals into low-level commands. Figure 5 shows our experimental *EEG*-based BCI system.

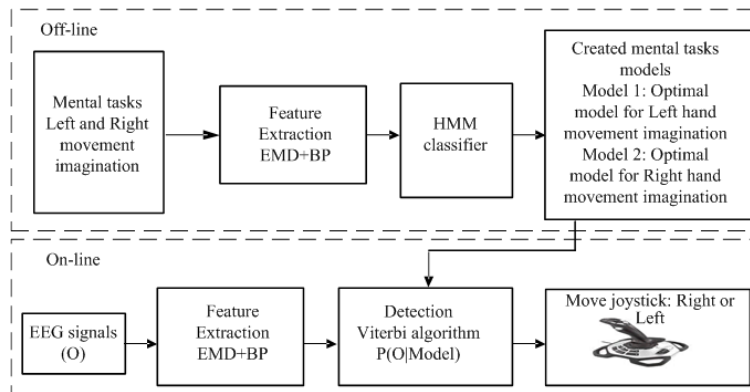


Figure 5. Experimental *EEG*-based BCI system used for the reinforcement of a joystick movement.

1) **Off-line phase:** In the first step, the *EEG* signals are recorded while subjects imagine a hand and left movement (see section II-A). The second step is the preprocessing and feature extraction of *EEG* data. In this step, we implemented our method to extract the relevant features of the *EEG*. This method is based on the combination of *EMD* and *BP* (see section II-B). In the third step, we implemented the *HMM* classifier to assign a model to each motor imagery task:  $\lambda_1$  for left hand movement imaginations and  $\lambda_2$  for right hand movement imaginations.

2) **On-line phase:** Once the motor imagery is identified by one of the two models, a low-level command can be then associated to this mental task. So, the on-line classification result is used to reinforce the movement of the joystick (right and left). This mechanism was implemented with Simulink/MathWork. *EMD* + *BP* and Viterbi algorithm (Rabiner, 1989) (see section II-C) are implemented as an embedded function in Simulink in order to identify the mental task on-line. Viterbi algorithm has two inputs data, the first input is the *EEG* data and the second is the two models already constructed in the first step (offline). The Viterbi-based recognition result is translated into a command to reinforce the movement of the joystick. As a function of the *EEG* data generation (left or right hand movement imagination), the movement of the joystick will be reinforced to left or to right consequently.

### 3. CONCLUSION

In this work, we were focused on two phases in the framework of motor imagery-based BCI systems. The first phase concerns the feature extraction and classification stages. A combination of a nonlinear feature extraction method based on the Empirical Mode Decomposition (*EMD*) and the conventional band power (*BP*) is proposed. This method consists first to apply the *EMD* on the raw *EEG* signals to obtain the *Intrinsic Mode Functions (IMFs)*. Then, the different *IMFs* are analyzed to keep only the active frequency band powers corresponding to  $\mu$  and  $\beta$  rhythms in BCI-related mental task *EEG* signals. Finally, the conventional *BP* of the active frequencies is applied to the relevant reconstructed signal. A nonlinear classifier based on *hidden Markov models (HMMs)* is employed to evaluate this feature extraction approach. The second phase is realized in on-line mode. In this phase, we translated the *EEG* signals corresponding to the imagination of left and hand movement to a low-level command. This system allows subjects who suffer from severe motor disabilities to move better a joystick to right and left. The proposed method seems promising and gives better preliminary classification results for three abled young subjects. In our future work, we will investigate the performance improvements by considering more number of subjects in order to obtain more significant statistical results and we will also generalize this method to other types of movements.

## REFERENCES

- Wolpaw, J. R., Birbaumer, N. J., McFarland D., Pfurtscheller, G. and Vaughan, T. M., 2002. Brain-computer interfaces for communication and Control. *In Clinical Neurophys*, Vol. 113, pp. 767-791.
- Wolpaw, J. R. and McFarland, D., 1994. Multichannel EEG-based brain-computer communication, *In Electroencephalogr Clin Neurophysiol*, Vol. 90, No. 6, pp. 444-449.
- Guger, C., Harkam, W., Hertnaes, C. and Pfurtscheller, G., 1999. Prosthetic control by an EEG-based brain-computer interface (BCI), *In Proceedings AAATE'99*, pp. 590-595.
- Mason, S. G. and Birch, G.E., 2003. A general framework for brain-computer interface design, *In IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 11, No. 1, pp. 70-85.
- Pfurtscheller, G., 1999. EEG Event-related Desynchronisation (ERD) and Event related Synchronisation (ERS). *In Niedermeyer*, pp. 958-967, E., Lopes da Silva, F.H. (Eds.), *Electroencephalography. Basic Principles, Clinical Applications, and Related Fields*, forth ed., Williams and Wilkins, Baltimore.
- Pfurtscheller G. and Neuper, C., 2001. Motor imagery and direct braincomputer communication. *In Proceedings of the IEEE*. Vol. 89, No. 7, pp.1123-1134.
- Herman, P., Prasad, G., McGinnity, T. M. and Coyle, D., 2008. Comparative analysis of spectral approaches to feature extraction for eeg-based motor imagery classification. *In IEEE Transactions on Neural Systems and Rehabilitation Engineering*, Vol. 16, No. 4, pp. 317-326.
- Penzel, T. and Conradt, R., 2000 .Computer based sleep recording and analysis. *In Sleep Medicine Reviews*, Vol. 4, No. 2, pp. 131-148.
- Lo, M., Tsai, P., Lin, P., Lin, C. and Hsin, Y. L. 2009. The Nonlinear and nonstationary Properties in EEG Signals: Probing the Complex Fluctuations by Hilbert-Huang Transform. *In Advances in Adaptive Data Analysis*, pp.461-482.
- Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N. C., Tung, C. C. and Liu, H. H.,1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *In Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, Vol. 454, No. 1971, Vol. 454, pp. 903-995.
- Rabiner, L. R., 1989. A tutorial on hidden Markov models and selected application in speech recognition. *Proceedings IEEE*, Vol. 77, No. 2, pp. 267-296.
- Niedermeyer, E., Lopes da Silva, F., 2004. *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. Lippincott Williams & Wilkins, ISBN 0781751268, 9780781751261.
- Gabor, D., 1946. Theory of communication. *Proceedings IEE*, vol. 93 (III), pp. 429- 457.
- Pfurtscheller, G., Neuper, C., Flotzinger, D. and Pregenzer, M., 1997. EEG-based discrimination between imagination of right and left hand movement. *Electroenceph clin Neurophysiol*, Vol. 103, No. 6, pp. 642-651.
- Welch, P. D., 1967. The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms. *In IEEE Transactions on Audio Electroacoustics*, Vol. AU-15, pp. 7073.
- Schwarz, G. , 1978.Estimating the dimension of a model," *Ann. Statist.*, Vol. 6, No. 2, pp. 461-464.
- Helmy, S., Al-ani, T., Hamam, Y. and El-madbouly, E., 2008. P300 Based Brain-Computer Interface Using Hidden Markov Models," *ISSNIP 2008*, Sydney, 15-18 Dec.
- Kbler, A., Mushahwar, V. K., Hochberg, L. R. and Donoghue, J. P., 2006. BCI meeting 2005-workshop on clinical issues and applications. *In IEEE Transactions on Neural Systems and Rehabilitation Engineering*. Vol. 14, No. 2, pp. 131-134, 2006.

# TOWARDS AN ONTOLOGY-BASED COMMUNITY OF AGENTS FOR PERSONALISATION OF SERVICES FOR DISABLED STUDENTS

Julius T. Nganji and Mike Brayshaw

*Department of Computer Science*

*University of Hull, Cottingham Road, Hull, HU6 7RX, United Kingdom*

## ABSTRACT

Disability legislations such as the Disability Discrimination Act 1995, 2005 and Americans with Disabilities Act 1990 call on service providers to make “reasonable adjustments” and “reasonable accommodation” respectively to their services in order to accommodate the needs of disabled people. Higher education institutions as service providers can respond through their services. Nevertheless, there is an increasing challenge in using digital media to deliver services due to the numerous problems associated with inaccessibility of some online systems with assistive/adaptive technologies. Personalisation is a solution to such problems as it provides content to students based on their needs. However, very little of such personalisation has been targeted towards disabled students and hence, this paper proposes an ontology approach for a semantic web community of agents that personalises services to disabled students in higher education. We present the architecture of such an agency, including a disability-aware semantic web agency development methodology and also present the ADOOLES ontology employed for such personalisation.

## KEYWORDS

Personalisation, services, higher education, disabled students, ADOOLES, ontology, society of agents.

## 1. INTRODUCTION

An increasing number of disabled students worldwide enter into higher education institutions every year. In the UK for instance, the Higher Education Statistics Agency (HESA) estimated the number of such students in the 2007-2008 academic year to be about 62,510 (HESA, 2009). This increasing number makes delivering education online more challenging as disabled students may have varying requirements based on their specific needs. To solve this problem, educational institutions utilise assistive technologies to assist disabled students in their learning. The problem however is that some of these technologies are not compatible with some digital learning systems, resulting in an exclusion of some disabled students if the institutions cannot handle the disability (Steyaert, 2005). The Disability Discrimination Act (DDA) was enacted in 1995 and extended in 2005 to prohibit discrimination against disabled people. Service providers are thus required to make “reasonable adjustments” to their services to meet the needs of disabled people. The Special Educational Needs and Disability Act 2001 (SENDA) introduces the right for disabled students to have equal opportunities to contribute and benefit from education and not to be discriminated against. Higher education institutions amongst other solutions have resorted to personalisation of e-learning and other services. The Americans with Disabilities (ADA) Act 1990 also prohibits discrimination against disabled people.

The problem of searching and retrieving online information continues because the current web is not meaningful and hence using search engines to retrieve information is in some cases a difficult task as the results are often numerous and sometimes irrelevant. Intelligent and semantically rich agents forming a community or society of agents (Minsky, 1986) which could empower the students based on their needs and requirements analyses are therefore needed. The semantic web offers a solution as it is by its nature more meaningful. Hence, researchers are currently using semantic web technologies to personalise learning and services (e.g. Brut and Braga, 2008; Razmerita and Lytras, 2008; Henze et al, 2004, Nganji et al, 2011) and to facilitate search and retrieval of information. However, as semantic web technologies are used for personalisation, little seems to be done to consider the needs of disabled people and to personalise services

for them thus inaccessibility problems continue, requiring more robust agents which understand their needs. This study therefore proposes a personalisation approach based on a disability ontology containing information on various disabilities, which can be used to present disabled people with services that are suitable for their specific needs.

This paper is organized as follows: section 2 presents a disability-aware approach to developing communities or societies of agents that are suitable for personalising services for disabled people. Section 3 presents the architecture of a semantic web agency for personalising services for disabled students in higher education, discussing the various agents and the roles they play in personalisation. The paper concludes with a summary of the contents of this work and future directions.

## 2. DISABILITY-AWARE SEMANTIC WEB AGENCY DEVELOPMENT

The education of disabled people needs to consider their needs, thus being disability-aware (Leicester, 2001). Similarly, in implementing agents or systems for disabled people in higher education, it is important to consider their specific needs, without which the system which should include them, would exclude them. This follows some procedures common to software engineering such as the various models for software development (Boehm, 1988; Royce, 1987). Nevertheless, those models would be insufficient without considering the disabled person by incorporating their needs at crucial stages. By incorporating the needs of disabled people at each stage of development, the product will be both accessible and usable. We propose a methodology to follow in implementing a disability-aware agency (Fig. 1). These are not separate stages that need to be completed before moving on, but constant review is needed for the agency to be improved during implementation.

Other researchers have considered design for people with special needs (e.g. Patrizia et al, 2009; Senatore et al, 2008). A disability-aware agency development begins with a consideration of the needs of the disabled user for which it is intended. Needs analysis is informed by literature, consultation of experts in the specific disability as well as potential users. Creating multi-agent personas further helps in having a clear idea of how the agency would be implemented, based on the needs of the user. These all inform the functional, accessibility, usability and maintenance needs of the agency. The developer will then need to acquire relevant technologies for implementation based on a review of related communities of agents and/or by consulting an expert in the field.

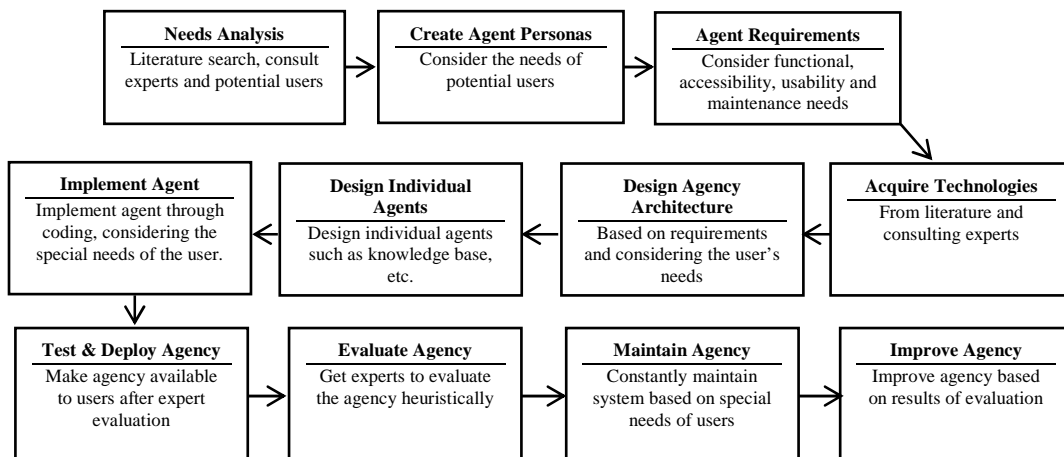


Figure 1. Disability-aware semantic web agency development methodology.

The agency architecture could then be designed, based on the requirements analysis and the needs of the user which will enable the developer to design various agents and thence to coding, testing, deploying and evaluating the agency, constantly maintaining it as need arises. Other researchers have developed a cognitive architecture based on the “Society of Mind” (e.g. Davis and Vijayakumar, 2010). We have developed an ontology-driven architecture for a semantic web agency based on a non-cognitively motivated “Society of Agents” metaphor; which personalises services for disabled students as presented in section 4.

### 3. AN ONTOLOGY FOR PERSONALISING SERVICES FOR DISABLED STUDENTS

Various disabilities exist amongst students in higher education. Some of these students with disabilities may require some form of assistance which could be provided through assistive technologies, human support workers or through animals such as guide dogs for blind people. These assistive mechanisms could be vital for the disabled students accessing various services in the university such as the Study Advice Service, Disability Services, Library, etc. Paradoxically, the use of such assistive technologies depends on the extent of some abilities. For instance, the use of screen readers and guide dogs may be recommended for students with very low ability to see. Using the screen reader also depends on the user's ability to hear as screen readers read out content to the user who depends on such form of communication.

Modeling an ontology to personalise services for disabled students could use the following concepts: *Abilities*, *Disabilities*, *Assistive Mechanisms* and *Services* which contain subclasses. Disability could have sub classes such as *Mental Disability*, *Physical Disability*, *Specific Learning Disability* and *Unseen Disabilities* as seen in Fig.2. Those subclasses are further divided into subclasses.

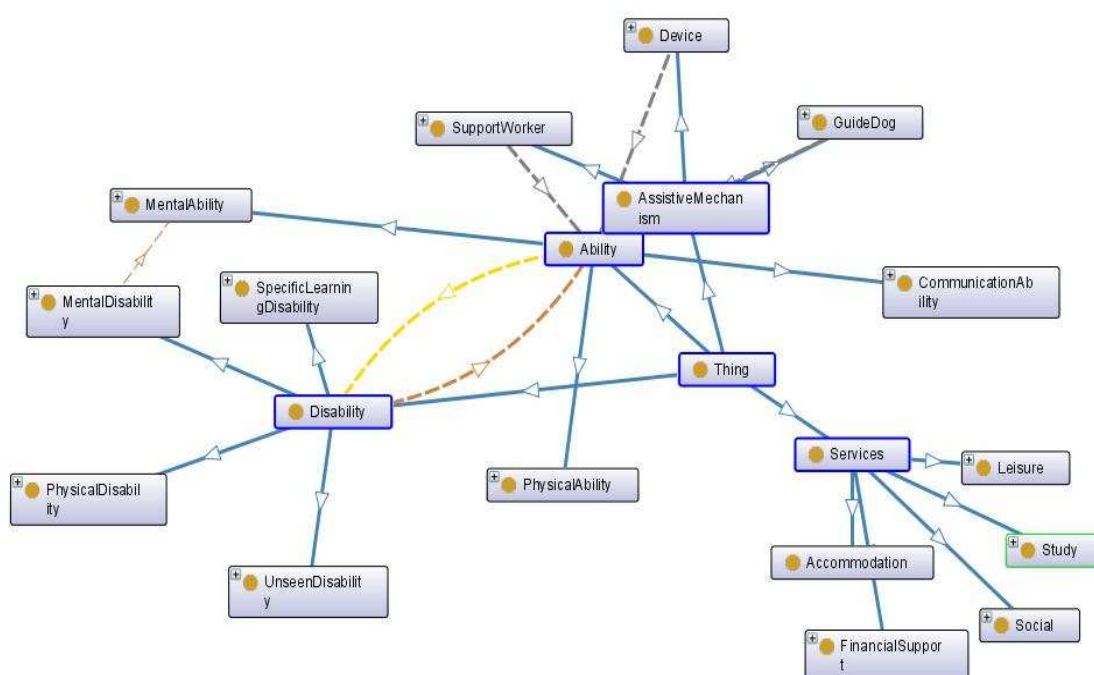


Figure 2. The ADOOLES ontology showing some main classes and subclasses

We have designed an ontology for personalising services for disabled students using OWL (Web Ontology Language), modeled with Protégé. Each concept of the ontology known as ADOOLES (Abilities and Disabilities Ontology for Online Learning and Services) is represented as an OWL class. ADOOLES incorporates some concepts, classes and properties contained in the ADOLENA ontology (Keet et al, 2008) but include the *Services* class which is not part of the ADOLENA ontology. Additionally, the classification of the ADOOLES ontology includes disabilities such as Specific Learning Disability which is not part of ADOLENA. The *Services* class of ADOOLES contains the following subclasses: *Accommodation*, *FinancialSupport*, *Leisure*, *Social* and *Study*. These are the services offered in some higher education institutions that are accessed by disabled students. By accessing these services regularly, disabled students could receive support that would be helpful to their study and overall experience at the university. Accommodation for a student with mobility difficulties such as those using wheelchairs would need to be very accessible such as on the ground floor or accessible through lifts and ramps. A dyslexic student might need support from the Study Advice Service who will provide resources in alternative formats that could help them in their study while the Disability Services might provide alternative examination arrangements. The



*AssistiveMechanism* class contains subclasses such as *Device*, *GuideDog* and *SupportWorker*, with some subclasses. A human support worker may be required in some instances to provide physical support to some disabled students in lecture rooms while assistive technology devices could be employed to provide support in accessing online information.

The OWL classes also have properties such as Object properties linking two or more individuals. For instance, in our ontology, the Object property *isSuitableFor* links a *LowerLimbMobility* individual with a *GroundFloorAccommodation* individual. Various other Object and DataType properties are used in the ontology. The ontology constitutes a key aspect of the Knowledge Representation Agent of our architecture as seen in Fig.3.

#### 4. ARCHITECTURE OF AGENCY FOR PERSONALISING SERVICES

The semantic web agency architecture has four main agents as depicted in Fig. 3. An agent is a basic unit or component within the community of agents which performs specific functions such as information retrieval, presentation, representation, etc.; while an agency constitutes communities of distributed agents dynamically corresponding to deliver an effective user solution. The agents work together to present the user with their requested service when they interact with the system from its user interface on an electronic device (computer, mobile phone, PDA). The agents and their functions are as follows.

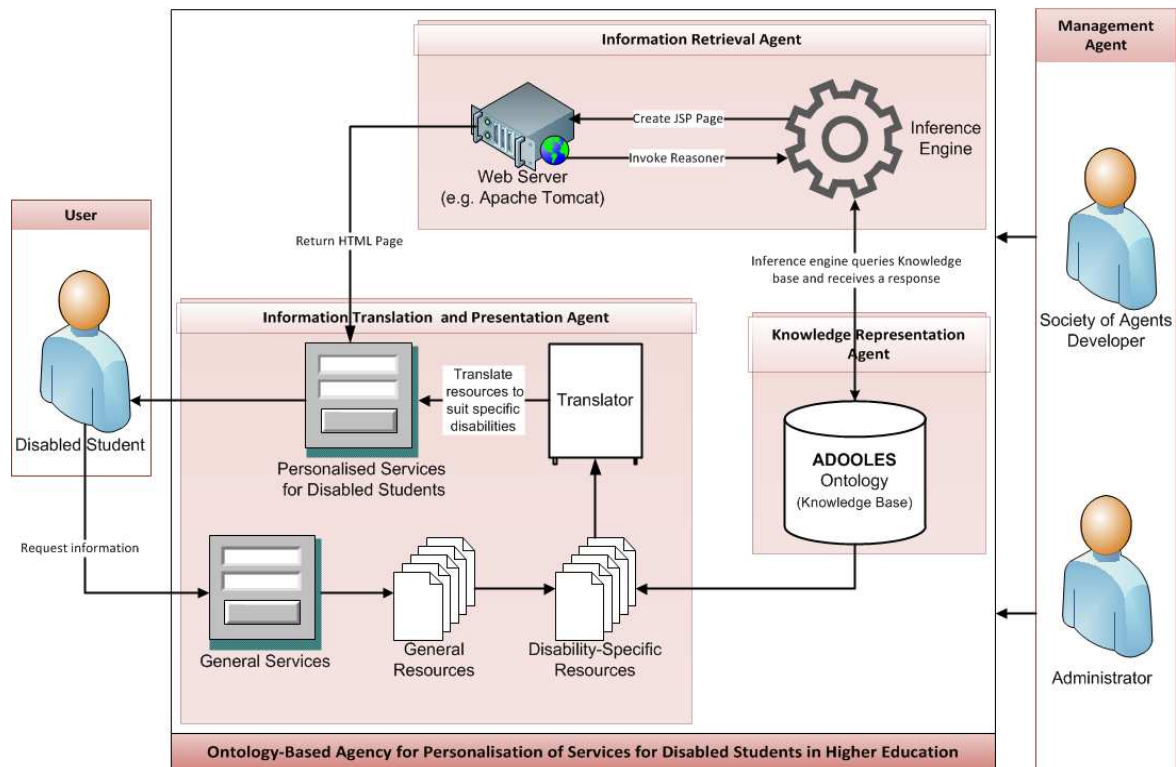


Figure 3. The semantic web community of agents for personalising services for disabled students in higher education is made up of the *information translation and presentation agent*, *knowledge representation agent*, *information retrieval agent* and the *management agent*

##### 4.1 Knowledge Representation Agent

This agent represents domain knowledge in a format that could be easily understood and interpreted by the agency. Its inference engine can easily query the knowledge base and obtain specific results which could be

used to personalise the services that have been requested. The agent also provides a means for assistive technologies to interact with it, ensuring compatibility and interoperability.

This agent has a disability ontology specified in OWL (Web Ontology Language). The ontology contains various concepts such as Ability, Disability, Assistive Mechanisms and Services, which all form the main classes. The ontology we have developed is known as Abilities and Disabilities Ontology for Online LEarning and Services (ADOOLES) and incorporates some concepts, classes and properties contained in the ADOLINA ontology (Keet et al, 2008) as already described in section 3.

## 4.2 Information Translation and Presentation Agent

This agent enables the disabled student to obtain content suitable for their specific needs. Initially, the student interacts with the semantic web society of agents through the user interface from their device, which first presents them with content comprised of several general services. The agency identifies the student as one with special needs through an authentication mechanism and retrieves the services that are specific and suitable for their disability. This is facilitated by the translation module which translates and presents the services to the student in an appropriate format. The translator for instance could accurately and fully describe an image to a blind user by semantically enriching it with comprehensive ontological descriptions which are then translated into text for interpretation by a screen reader, thus facilitating understanding of any visual resource as represented in Fig. 4.

Initially, an assistive technology such as a screen reader may be unable to describe the image to the blind user because it is not semantically enriched with accurate descriptions of the image and in some instances, the web designer or developer may not have adhered to Web standards for accessibility. There is therefore a semantic description barrier which needs to be overcome by semantically annotating the image with comprehensive descriptions of its content through automatic processes. This image can then be integrated into the presentation agent for interpretation by the assistive technology which conveys the meaning of the image to the user for better understanding.

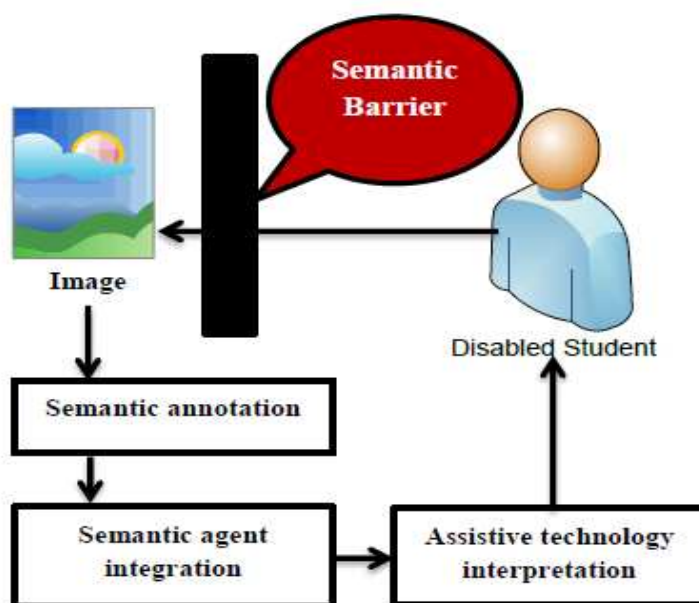


Figure 4. Image interpretation in the translator module

## 4.3 Information Retrieval Agent

This agent is one of the key agents of the society of agents which interacts with the knowledge representation agent to collect information on the services that are being requested, transmitting it to the information

translation and presentation agent. It has an inbuilt inference engine which queries the ontology and a web server (e.g. Apache Tomcat) interacting with the inference engine to present the information to the user.

#### **4.4 Management Agent**

This agent has a visual interface which enables administrators such as lecturers, agency administrators and developers to manage the student's accounts and to solve some problems they may encounter whilst interacting with the society of agents.

### **5. RELATED WORK**

Ontologies have been successfully employed to help provide effective solutions to the difficulties faced by disabled people. Coyle and Doherty (2008) for instance described the potential of using ontologies in developing interactive systems that can support mental health interventions. Lorenz and Horstmann (2004) on the other hand used the Resource Description Framework (RDF) to facilitate access to graphically represented information for blind people by semantically annotating the graphical information with RDF.

One of the earliest adaptation models is AHAM (De Bra et al, 1999), Adaptive Hypermedia Application Model which is an extension of the Dexter model. In this model, the topmost layer, the run-time layer represents the user interface on top of a presentation specifications layer and a storage layer containing three models: an adaptation model which combines elements from a domain model and user model to describe an event, a domain model which contains a conceptual representation of the application domain and a user model containing conceptual representations of all aspects of the user relevant for the adaptive hypermedia application. Whilst the user model represents aspects of the user, it does not specifically deal with disabled users, and does not show how such model can be employed to personalise or adapt services for disabled people.

The FOHM (Fundamental Open Hypermedia Model) model (Millard et al, 2000) defines a common data model and set of related operations that are applicable for the three hypertext domains of navigational hypertext, spatial hypertext and taxonomical hypertext.

With the introduction of the Semantic Web, newer architectures and systems such as the Personal Reader (Henze and Herrlich, 2004) make use of Semantic Web technologies for personalisation of services. In this framework, different web services cooperate with each other by exchanging RDF documents to form a Personal Reader instance. One of the key advantages with this framework is that it enables users to select services which provides extended functionality such as personalization services, combining them into a Personal Reader instance. This framework as others is not specific for disabled people and does not describe how to deal with specific needs.

There is evidently a great gap in using ontologies to personalise services to disabled people, which necessitates an adaptation of existing architectures to provide such personalisation or newer architectures. Our architecture presented in this work although specific for disabled people, can be used for non-disabled people, thus being inclusive.

### **6. IMPLEMENTATION AND EVALUATION**

With the disability ontology already developed, the semantic web society of agents will now be fully realized following the procedures described herein. The prototype is currently being implemented using Java Server Pages and Servlets with the translation agent implemented in Java. This makes use of a text-to-speech interface for translating information for those who request it. Apache Tomcat 7.0 is used as the server and NetBeans IDE 6.9.1 provides the platform for development. The prototype will tackle the different disabilities by providing specific personalisation based on disabilities. The ontology is employed to provide automatic recognition of a disabled user and to personalise services to them based on their specific needs.

Once the prototype is fully implemented, it will be evaluated heuristically against some established heuristics (Squires and Preece, 1999), thus addressing usability issues that may arise with the interface. By

following accessibility and usability guidelines incorporated in the disability-aware semantic agency development methodology, problems of accessibility and usability will be greatly minimised. The heuristic evaluation will also give an indication of the strengths and weaknesses as well as the applicability of the ontology-based approach for personalising services for disabled people (Longpradit et al, 2008).

## 7. CONCLUSION AND FUTURE WORK

In this paper, we have presented a conceptual architecture for a semantic web community of agents which personalises services for disabled students in higher education. The agency is made up of four main agents which work together to present the user with the services they request, based on their specific needs as identified by the system. The methodology for developing such agency has been presented which considers the needs of disabled students at crucial stages of development.

Future work will assess the success of the prototype in enhancing study for students with disabilities by assessing their performances pre and post interaction with the prototype over time, compared to other students with no such access.

## ACKNOWLEDGEMENT

The authors wish to thank the Department of Computer Science, University of Hull for a scholarship that enabled this research to be conducted. We also thank the anonymous reviewers for their insightful comments.

## REFERENCES

- Boehm, B.W. 1988. A Spiral Model of Software Development and Enhancement. *ACM SIGSOFT Software Engineering Notes*, Vol. 11, Issue 4, pp14-24.
- Brut, M. and Buraga, S.C. 2008. An ontology-based approach for modelling grid services in the context of e-learning. *International Journal of Web and Grid Services*. vol. 4, issue 4, pp. 379-394.
- Coyle, D. and Doherty, G. 2008. Towards Ontologies for Technology in Mental Health Interventions. *Proceedings of the 2008 First International Workshop on Ontologies in Interactive Systems*, 2008 Washington, DC, USA. IEEE Computer Society.
- Davis, D.N. and Vijayakumar, M.V. 2010. A "Society of Mind" Cognitive Architecture Based on the Principles of Artificial Economics. *International Journal of Artificial Life Research (IJALR)*, Vol.1, Issue 1, pp.51-71.
- De Bra, P. et al, 1999. AHAM: A Dexter-based Reference Model for Adaptive Hypermedia. *Proceedings of the ACM Conference on Hypertext and Hypermedia*.
- Henze, N. and Herrlich, M. 2004. The Personal Reader: A Framework for Enabling Personalization Services on the Semantic Web. *In Proceedings of the Twelfth GI Workshop on Adaptation and User Modeling in Interactive Systems (ABIS 04)*, Berlin, Germany.
- Henze, N. et al. 2004. Reasoning and Ontologies for Personalized E-Learning in the Semantic Web. *Educational Technology & Society*, 7 (4), pp. 82-97.
- HESA, 2009. Students and Qualifiers Data Tables. Higher Education Statistics Agency. Available: [http://www.hesa.ac.uk/index.php?option=com\\_datatables&Itemid=121&task=show\\_category&catdex=3](http://www.hesa.ac.uk/index.php?option=com_datatables&Itemid=121&task=show_category&catdex=3), [Accessed: 31 October 2009].
- Keet, C.M. et al, 2008. Enhancing web portals with Ontology-Based Data Access: the case study of South Africa's Accessibility Portal for people with disabilities. *Fifth International Workshop OWL: Experiences and Directions (OWLED'08)* Karlsruhe, Germany.
- Leicester, M. 2001. A moral education in an ethical system. *Journal of Moral Education*, vol. 30, issue 2, pp.251-260.
- Longpradit, P. et al, 2008. An Inquiry-led Personalised Navigation System (IPNS) using multi-dimensional linkbases. *New Review of Hypermedia and Multimedia*, 14 (1). pp. 33-55.
- Lorenz, M. and Horstmann, M. 2004. Semantic Access to Graphical Web Resources for Blind Users. *3rd International Semantic Web Conference (ISWC 2004)*, Hiroshima, Japan.

- Millard, D., et al, 2000. FOHM: A Fundamental Open Hypertext Model for Investigating Interoperability between Hypertext Domains. In: *Proceedings of the Eleventh ACM Conference on Hypertext and Hypermedia HT'00*, pp. 93-102.
- Minsky, M. 1986. *The Society of Mind*. Simon and Schuster, New York.
- Nganji, J. et al. 2011. Ontology-Based E-Learning Personalisation for Disabled Students in Higher Education. *ITALICS*, February 2011, vol. 10, issue 1, pp.1-11.
- Patrizia, M. et al, 2009. A Robotic Toy for Children with special needs: From requirements to Design. *Proceedings of the 11th International Conference on Rehabilitation Robotics*, Kyoto, Japan, pp. 1070-10-75.
- Razmerita, L. and Lytras, M.D. 2008. Ontology-based user modeling personalization: Analyzing the requirements of a Semantic Learning Portal. *Proceedings of the 1st World Summit on the Knowledge Society (WSKS 2008)*, Rhodes, Greece. Vol. 5288, pp. 354-363.
- Royce, W. W. 1987. Managing the Development of Large Software Systems, *Proc. 9th. Intern. Conf. Software Engineering*, IEEE Computer Society, pp. 328-338.
- Senatore, F. et al, 2008. Development of a Generic Assistive Platform to Aid Patients with Motor Disabilities. *Proceedings of the 14th Nordic-Baltic on Biomedical Engineering and Medical Physics*, Riga, Latvia, pp. 168-171.
- Squires, D. and Preece, J. 1999. Predicting quality in educational software: Evaluating for learning, usability and the synergy between them. *Interacting with Computers*, 11: 467-483.
- Steyaert, J. 2005. Web-based higher education: the inclusion/exclusion paradox. *Journal of Technology in Human Services*, vol. 23, issue 1, pp. 67-78.

# SUPPORTING DOWN PEOPLE IN CUTTING OPERATIONS THROUGH HAPTIC TECHNOLOGY

Mario Covarrubias, Monica Bordegoni and Umberto Cugini

*Mechanical Department, Politecnico di Milano  
Via La Masa 1, 20156, Milan, Italy*

## ABSTRACT

The paper deals with the application of haptic technology when improving manual skills of people with specific disorders, such as Down syndrome, mental retardation, etc. The development of a cutting haptic system is the base of this paper to show specificity of the development and verification of hardware for the given group of handicapped users. The paper shows suitability of using haptic technology for a concrete application, which is designed for user with specific disorders. The cutting operation is performed using a hot wire tool, which is linked through an R-R mechanism to a PHANTOM device. The haptic cutting device is able to cut soft materials as expanded and extruded polystyrene foam from 5 to 20 mm thickness. The device is driven under the user's hand movement and assisted through the Magnetic Geometry Effect (MGE). The haptic cutting device has been used as an input system for tracking the sketching movements made by the user according to the visual feedback received from a physical template without haptic assistance. Then the cutting device has been used as an output system that provides force feedback capabilities. Finally, the system performance has been evaluated by comparing the analysis of the tracking results with the final polystyrene components.

## KEYWORDS

Haptic technology, assisted cutting system, unskilled people.

## 1. INTRODUCTION

There is a potential group of users with specific disorders who could make use of the development of haptic technology. E.g. people with Down syndrome, forms of mental retardation and other development defects, which cause problems related to precision of movement, coordination of force, speed and reduced efficiency in performance as compared with healthy people. This modern technology can play a role of important supporting and compensational means for a group of users with specific needs. Down syndrome people differ from the normal one in their age in various dimensions for instance in terms of cognitive, language input and usage, physical abilities and motor as well as social and individual characteristics.

In case of developmental disorders, such as Down Syndrome, motor control may be greatly affected [Blank et al., 1999]. People with Down Syndrome may demonstrate reduced sensory acuity, lengthened reaction time and altered postural responses to perturbation. One of the hypotheses suggests that the source of motor difficulties originates in deficit of the central representation of actions. As suggested by the literature [Blank et al., 1999, Kurillo et al., 2005] practice can have positive influence on the motor skill. Haptic guidance has a significant value in many applications, such as medical training [Liu et al., 2003], hand writing learning [Teo et al., 2002], and in applications requiring precise manipulations [Ahlström, D., 2005]. The guiding force can be either generated by a force-feedback device, such as a Phantom haptic device [PHANToM, 2011]. There are, however, only very few assisted applications for unskilled people in order to promoting them in a specific employment role. The aim of our study was to assess how the sketching control movements under haptic feedback are affected in people with Down syndrome. 2D cutting operations are directly related with sketching operations. Sketching is one of the most complex human activities in which the hand movements are controlled by the central nervous system, which regulates the activity of the hand and arm muscles to act in synergy. The central nervous system receives dynamic feedback information from visual sensors and from other body sensors located on the skin, muscles and joints while regulating the motor

output. Many children with Down syndrome who have received family support, enrichment therapies, and tutoring have been known to graduate from high school and college, and enjoy employment in the work force. It is well known that everyday individuals with Down syndrome strive to accomplish the same goals as everyone else. They want self-fulfillment and inclusion into the communities' accomplishments and activities. On a daily basis people with Down syndrome venture out into their communities to jobs and schools and strive to reach their full potential. There are many companies today that agree to employ persons with Downs in specific parts of their company. There are other places where the entire campus employees people with Down syndrome. Some places even have bus service where they come and pick up the people as well as bring them home. We have designed the cutting haptic device for providing assistance movements while cutting through the haptic point-based approach.

## 2. RELATED WORK

The research concerning haptic technology has increased rapidly during the last few years, and results have shown the significant role that haptic feedback plays principally in industrial design. Designers are beginning to realize the advantages of haptic displays in order to help individuals to overcome the challenges experienced when accessing and exploring the web [Kuber 2007]. The touch modality has also been shown to make it possible for visually impaired to explore and navigate virtual environments [Sjostrom 2001].

The interaction is enriched by the use of the sense of touch, since visually impaired users can identify objects and perceive their shape and texture. There are, however, only very few assisted applications for unskilled people in order to promote them in a specific employment role. Force feedback guidance using the haptic point-based approach is deemed to be a good method for facilitating 2D cutting operation in unskilled people in general and Down and blind people in particular. Haptic interface technology enables individuals who are blind to expand their knowledge by using an artificially made reality built on haptic and audio feedback. Research on the implementation of haptic technologies within Virtual Environments (VEs) has reported the potential for supporting the development of cognitive models of navigation and spatial knowledge with sighted people (Witmer et al., 1996; Giess et al., 1998; Gorman et al., 1998; Darken and Peterson, 2002) and with people who are blind as well (Colwell et al., 1998; Jacobson et al., 1998).

## 3. HAPTIC CUTTING DEVICE

The haptic cutting device concept has been realized in a series of virtual and physical prototypes to enable evaluation of its potential for improving 2D cutting operations. In order to enable the 2D cutting operations by the user, several configurations have been analyzed. This section describes the RR configuration.

### 3.1 The Concept

Figure 1 provides the isometric view of the CAD concept using the RR mechanism (2). The stylus (5) of the Phantom desktop is driven under the operator's movement and assisted by the Magnetic Geometry Effect (MGE). When this option is activated, a spring force tries to pull the sphere of the stylus (5) of the haptic device towards the surface of geometry. In fact, this effect is used in order to assist the user's hand.

The lengths of the links in the Phantom device determine the mechanisms kinematic properties such as workspace and manipulability. In order to find link lengths (subassembly 2), we have performed some analyses that display the kinematic properties.

Since the multifunctional device contains closed kinematic loops, it is necessary to consider over-constraint problems caused by joint axes that are misaligned during assembly or manufacture. We use carefully selected bearing configurations to avoid the over-constrain problem.

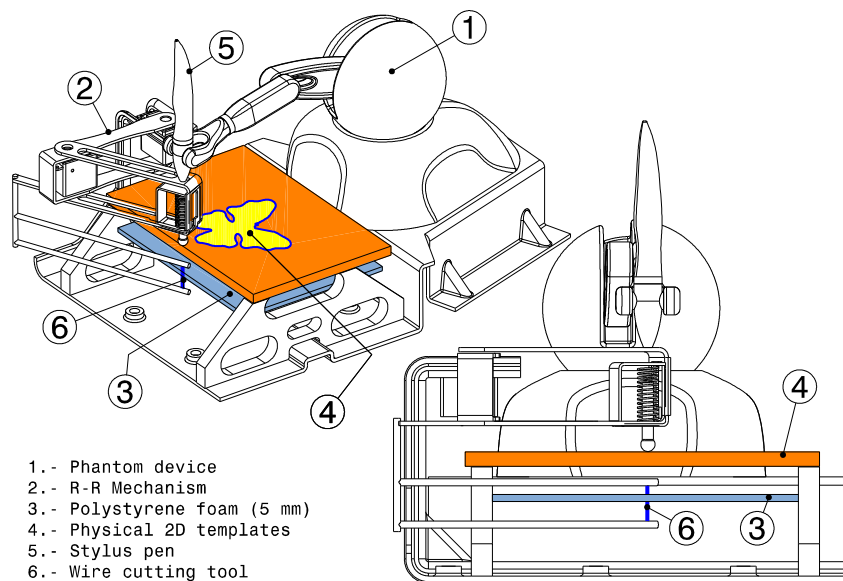


Figure 1. Haptic Cutting System.

For instance, pure revolute behavior is required, in order to avoid off-axis wobble. We use a single bearing at joints R1, R2 and R3. This allows enough off-axis wobble to resolve the over-constraint. In order to realize the proposed mechanism, we used the miniature bearings (7mm(OD) / 4mm(ID) and 5mm(OD) / 2mm(ID)) and precision-machined linkages. Link collisions have been carefully considered.

While user follow the 2D template by using the haptic guidance, the wire tool (6) cuts the polystyrene foam (3). The polystyrene foam is an interchangeable element.

#### 4. USE OF THE SYSTEM

As mentioned in the introduction, Down people is the target group of final user for the multifunctional device. Taking this into account, the mental load required to use the device should be reduced by means of using a co-located system [Jansson and Öström, 2004]. The term co-location has been used to describe a haptic and a visual display that is calibrated such that the visual and haptic coordinate systems are coincident. This means that the user can visually perceive an object in the same position in space as the haptic simulation [Wall et al. 2002].

This criterion simplifies the perception process required from users and facilitates its natural integration to allow fast reflexive motor responses to the haptic stimuli. In our device, this is achieved by using a 2D printed sketch aligned with the coordinate system of the 2D haptic sketch. The haptic rendering and the control of the haptic device are based on the H3D API platform (an open source API). Therefore, the geometric model of the virtual sketch and all the necessary physical properties are defined through a configuration file, written in the "X3D" format.

#### 5. ACCURACY VALIDATION

We have carried out several preliminary tests in order to validate the haptic cutting device while performing 2D tasks. The operations have been performed by tracking the stylus of the Phantom device through the DeviceLog command provided by the H3D API platform. The tracked sample rate has been used at 25 Hz.



## 5.1 Tracking Sketching Operations

In this task, a 2D printed sketch has been provided with the same coordinate system of the 2D haptic sketch. First, we ask for a free sketching by using the device without the haptic support, and then the same sketch has been designed by enabling the 2D haptic sketch as a guide path. The left column on Figure 2 shows the resulting path tracked by the user's sketching operation in which the haptic feedback is off.

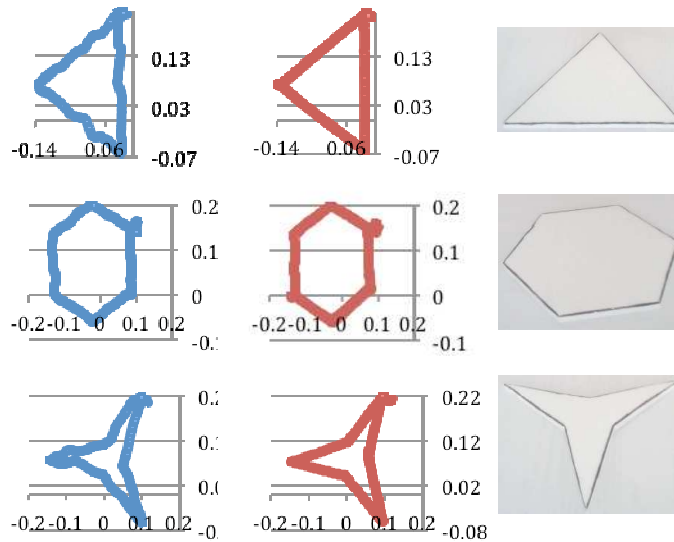


Figure 2. Sketching paths

The central column on Figure 2 shows the same sketch operation with the haptic feedback enabled; we can observe a strong difference in the accuracy of the operation. Similar task has been requested using basic geometries as a circle, rectangle, triangle, hexagon, and a star as can be seen from Figure 2. The most evident advantage provided by the 2D haptic sketches is the accuracy. The right column on Figure 2 shows the real polystyrene components (5 mm thickness). We have painted the edges in black in order to compare the quality of the polystyrene products. The sketching paths have been used in the assessment of the mechanical behavior of the real prototype.

## 5.2 Cutting Operation and 3D Assembly

In this preliminary task, several components have been manufactured by using the haptic cutting system. Figure 3 shows the ten polystyrene components required in the house assembly. The house product has been designed enabling both easy assembly and disassembly operations. In fact, all the components are symmetric and we have integrated several slots, which are used for the correct component's position in the assembly.

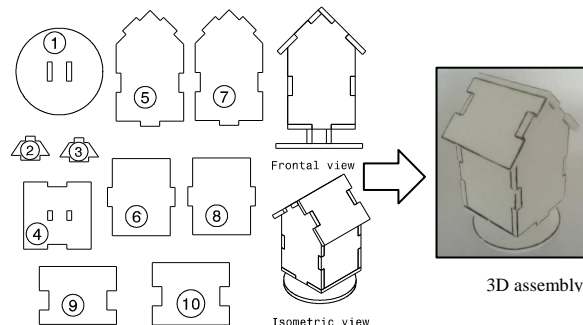


Figure 3. Cutting and assembly operations

## 6. PHYSICAL PROTOTYPE

The main structure used for the haptic cutting system has been designed taking into account some important considerations related to the use of sheet metal and aluminum components that implies: low inertia, light weight parts and low friction. Figure 4 shows the prototype of the haptic cutting system.

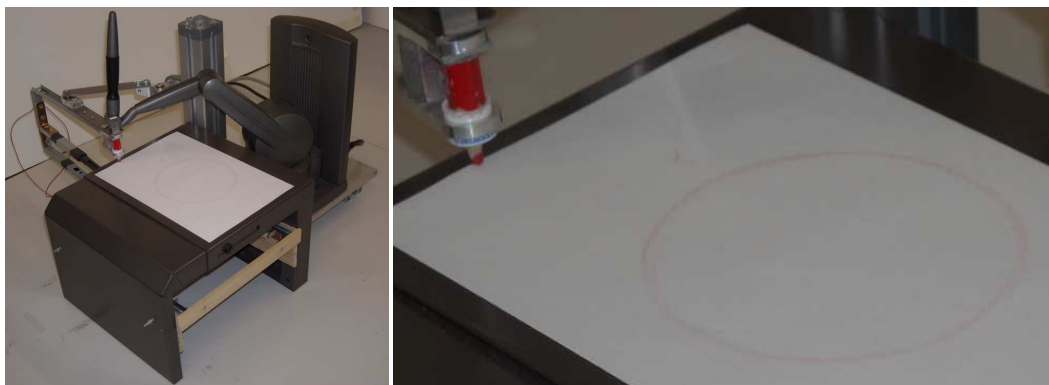


Figure 4. Haptic cutting system

In the manufacturing process some considerations have been made: the selection of revolute joints and its stiffness or resistance to all undesired motion, the shafts diameter, clearances and tolerances and mounting configurations of the components. Also, the mounting arrangement of the main structure has been designed to accommodate manufacturing tolerances. A 4 volts dc battery is connected to the cutting wire tool through the red wire. The 4 volts dc battery is required in order to safely produce the necessary heat to cut the polystyrene. The safety is really an important constraint that has been addressed in order to prevent any unsecured condition by the user. We decided to use a dc low voltage instead of an ac high voltage for producing the heat. In fact, 4 volts are enough for heating the wire used as cutting tool. In the same figure is also displayed the way in which the stylus of the Phantom device is linked to the RR mechanism.

## 7. PRELIMINARY TEST

We have performed a preliminary test with 6 users. Participants were down people; right-handed volunteers (5 females, 1 male) aged 8 to 28 years. Verbal instructions and familiarization period were issued. Table 1 shows the results in terms of radius (mm) while sketching a circle. R1 is the target circle; R2 and R3 are the internal and external circles respectively as can be seen in Figure 5.

Table 1. Sketch results, in terms of radius (mm)

<i>User</i>	<i>R1</i>	<i>R2</i>	<i>R3</i>
A	100.0	61	98
B	100.0	72	107
C	100.0	68	110
D	100.0	85	106
E	100.0	79	115
F	100.0	86	116

All participants were asked to perform a task that involved using a combination of visual and haptic functions in order to design a circle with 100 mm of radius.

The majority of participants were able to complete the specified tasks; the time taken to complete the task varied considerably between participants.

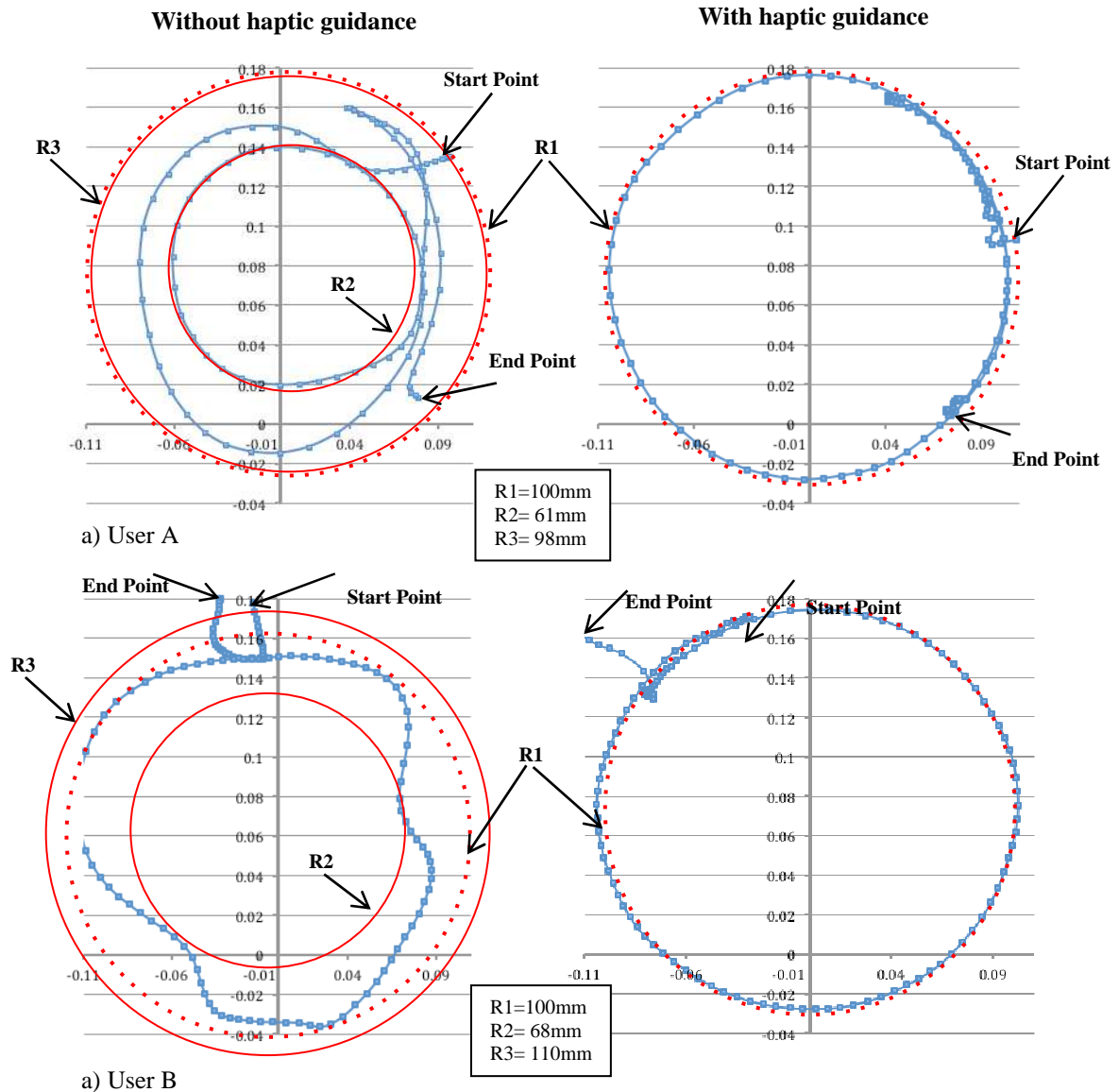


Figure 5. Two samples of the preliminary test

On the left is displayed the tracked motion without haptic guidance while sketching the circle. On the right is displayed the tracked motion with the haptic guidance. When comparing the final results on each of the six users, we can conclude that haptic guidance can benefit the performance in sketching tasks.

In summary, participants were generally pleased with the ease with which they were able to draw an almost “perfect” circle. The overall response to the experience was positive.

## 8. BENEFITS OF USING NEW TECHNOLOGY

A few researches have concluded that technology is very important to students with Down syndrome. It can also improve the effectiveness and the speed of the learning processes [Galanouli, et al. 2004; Andrews et al. 2007]. Several advantages have been suggested by [Black and Wood, 2003].

- Improving motivation: The learning experience is enhanced with sketches, which may increase interest and attention. People with DS are “visual learners” who learn best when information is presented visually.

- Immediate feedback: Users with DS are rewarded for their successes immediately, e.g. by creating a sketch or by cutting some funny objects. The haptic system never gets impatient or frustrated by repeated errors. Feedback is non-threatening and non-judgmental.

- Opportunities for practice: Users with DS need much more practice to acquire new skills and the haptic device can provide as many opportunities as necessary to repeat the same objective in exactly the same way allowing to proceed as fast or as slow as he or she wishes; the haptic system will “wait” for the child to respond without prompting them before they have time to fully process the information and construct their response.

## 9. DISCUSSION

The results of our study showed that the cutting haptic device help people during cutting operations by means of using haptic technology. However, we cannot compare the accuracy for example of the cutting components manufactured with the cutting haptic device with the ones manufactured with traditional machining processes or with CNC technology in which the components can be made with a high degree of accuracy. Nevertheless, with the haptic cutting system is also possible to design and create aesthetic and funny 2D sketches in which high precision is not required. The products made up with the cutting haptic device are hand-made assisted ones. The preliminary analysis process is implemented to identify research problems and to know the specifications of the cutting haptic system needs. The opportunity to create a haptic system that would make real difference in Down people’s life appeared to be a highly motivating factor. We are currently performing an evaluation with Down people in order to measure their learning improvements in 2D operations skills.

## 10. CONCLUSIONS

The paper proposes a cutting haptic device based on the point-based haptic approach. The main application of the haptic device is related to assisting unskilled people in the assessment and training of hand movements while performing cutting tasks. We explored user performance while using the haptic point-based approach. Results show that the effect of using the haptic cutting system increases the accuracy in the tasks operations. We can resume that the system leads to the satisfaction of the following objectives:

1. The coherence and collocation between the haptic and the physical 2D template is assured;
2. The force feedback as an additional perceptual channel for enhancing the interaction between user and computer;
3. The use of haptic cutting device for enhancing human skills and/or as a rehabilitation tool for disable people.

Further research, however, is still needed to improve the performance of the cutting haptic device by increasing the working volume.

## ACKNOWLEDGEMENT

The authors would like to thank Emanuele Fedeli for his support in the preliminary evaluation of the haptic cutting device. Emanuele Fedeli provides specialized manual training for Down people at Laboratorio Artimedia (Calolziocorte, Italy).

## REFERENCES

- Ahlström, D., 2005. Modeling and improving selection in cascading pull-down menus using fitts’ law, the steering law and force fields. In Proceedings of the SIGCHI conference on Human factors in computing systems, CHI ’05, ACM, pp. 61–70.

- Andrews, R., Freeman, A., Hou, D., McGuinn, N., Robinson, A., Zhu, J., 2007. The effectiveness of information and communication technology on the learning of written English for 5 To 16-year-olds, *British Journal of Educational Technology*, 2007, 38(2):pp.325-336.
- Black, B., Wood, A., 2003. Utilising information communication technology to assist the education of individuals with Down syndrome, *Down Syndrome Issues and Information*.
- Blank, R., Heizer, W., von Voss, H., 1999. Externally guided control of static grip forces by visual feedback age and task effects in 3-6 year old children and in adults. In *Neuroscience Letters*, volume 271, pages 41–44.
- Colwell, C., Petrie, H., Kornbrot, D., 1998. Haptic virtual reality for blind computer users. Paper presented at the Assets '98 Conference, Los Angeles, CA. Available in: <http://phoenix.herts.ac.uk/sdru/pubs/VE/colwell.html>.
- Darken, R.P., Peterson, B., 2002. Spatial orientation, wayfinding, and representation. In: Stanney, K.M. (Ed.), *Handbook of Virtual Environments Design, Implementation and Applications*. Lawrence Erlbaum Associates, Inc., New Jersey, pp. 493–518.
- Galanouli, D., Murphy, C., Gardner, J., 2004. Teachers' perceptions of the effectiveness of ICT-competence training, *Computers and Education*, 2004, 43 (1-2 SPEC ISS.), pp. 63-79.
- Giess, C., Evers, H., Meinzer, H.P., 1998. Haptic volume rendering in different scenarios of surgical planning. Paper presented at the Third PHANTOM Users Group Workshop, M.I.T., Massachusetts.
- Gorman, P.J., Lieser, J.D., Murray, W.B., Haluck, R.S., Krummel, T.M., 1998. Assessment and validation of force feedback virtual reality based surgical simulator. Paper presented at the Third PHANTOM Users Group Workshop, M.I.T., Massachusetts.
- Jacobson, R.D., Kitchin, R., Garling, T., Golledge, R., Blades, M., 1998. Learning a complex urban route without sight: comparing naturalistic versus laboratory measures. Paper presented at the International Conference of the cognitive Science Society of Ireland. Mind III, Ireland: University College, Dublin, Ireland.
- Jansson, G., and Öström, M., 2004. The effects of co-location of visual and haptic space on judgments of form. In *Proceedings of EuroHaptics 2004*, p.516.
- Kuber, R., Yu, W., And Mcallister, G. 2007. Towards Developing Assistive Haptic Feedback for Visually Impaired Internet Users, In *Proceedings of the 2007 Conference on Human Factors in Computing Systems*, San Jose, California, USA, 1525 – 1534.
- Kurillo, G., Gregorič, M., Goljar, N., Bajd, T., 2005. Grip force tracking system for assessment and rehabilitation of hand function. *Technol. Health Care*, 13, July, pp. 137–149.
- Liu, A., Tendick, F., Cleary, K., Kaufmann, C., 2003. A survey of surgical simulation: applications, technology, and education. *Presence: Teleoper. Virtual Environ.*, 12, December, pp. 599–614.
- PHANTOM device, SenSable Technologies Inc., <http://www.sensable.com>, accessed 16 February 2011.
- Sjostrom, C. 2001. Designing Haptic Computer Interfaces for Blind People. *Proc. ISSPA*, 1, 68-71.
- Teo, C., Burdet, E., and Lim, H., 2002. A robotic teacher of chinese handwriting. In *Proceedings of the 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, HAPTICS '02*, IEEE Computer Society.
- Wall, S., Paynter, K., Shillito, A., Wright, M., and Scali, S., 2002. The effect of haptic feedback and stereo graphics in a 3d target acquisition task. In *Proceedings of Eurohaptics 2002*, University of Edinburgh, 8-10th July, pp. 23– 29.
- Witmer, B.G., Bailey, J.H., Knerr, B.W., Parsons, K.C., 1996. Virtual spaces and real world places: transfer of route knowledge. *International Journal of Human-Computer Studies* 45, 413–428.

# INTERPRETATION OF AMBIGUOUS IMAGES INSPECTED BY THE STICKGRIP DEVICE

Tatiana V. Evreinova, Grigori Evreinov and Roope Raisamo  
*School of Information Sciences, University of Tampere*  
*FIN-33014 Tampere, Finland*

## ABSTRACT

Interpretation of ambiguous images perceived visually and relying on supplementary information coordinated with pictorial cues was selected to evaluate the usefulness of the StickGrip device. The ambiguous visual models were achromatic images composed from only two overlapping ellipses with various brightness gradients and relative position of the components. Inspection of images by the tablet pen enhanced with the pencil-like visual pointer decreased discrepancy between their actual interpretation and expected decision by only about 2.6 for concave and by about 1.3 for convex models. Interpretation of the convex images ambiguous with their inverted concave counterparts inspected by the StickGrip device achieved three times less discrepancy between decisions made and expected. Interpretation of the concave images versus inverted convex counterparts was five times more accurate with the use of the StickGrip device. We conclude that the kinesthetic and proprioceptive cues delivered by the StickGrip device had a positive influence on the decision-making under ambiguous conditions.

## KEYWORDS

Pictorial depth, surface computing, pen-based interaction, kinesthetic feedback, StickGrip haptic device.

## 1. INTRODUCTION

To perceive real world objects and events, the observers rely on pictorial cues and perceptual knowledge obtained through practice. The depth cues inform an observer about the surrounding objects, their size, height, shape, orientation and relative location. Knowledge about visual depth is inherently multimodal in nature. Since an observation is accompanied by the tactile, proprioceptive and kinesthetic senses in human fingers, in addition to pictorial cues, the visual system integrates sensory information gained from active manipulation by physical objects located in peripersonal space.

However, different modalities have evolved to complement the afferent information and not to duplicate each other. Even visual cues of depth can compete with each other, and with afferent information in another modality (Ernst & Banks, 2002; Dresch et al., 2002). Faconti and colleagues indicated that haptic information could confirm or reject the expectations inferred from visual mental model (Faconti et al., 2000). In some situations, the objective cues might be suppressed or superimposed being weighted based on prior knowledge (Boyd, 2001). Several cues can mislead an observer when they are inconsistent and presented together. Therefore, visual and haptic cues can create illusions which can influence the interaction and interpretation of pictures and alter the haptic experience (van Mensvoort, 2009; Faconti et al., 2000).

In recent research, Wijntjes with co-authors (2009) experimented with 3D virtual model of the ellipsoid, which could be viewed and touched at the same time. They have demonstrated that visual and haptic senses can supplement each other. Nevertheless, the subjects underrated the depth of relief relying only on visual perception, while haptic information did reduce visual ambiguity discovered. The authors concluded that the visual and manual investigation of 3D shape could improve the accuracy of depth perception in comparison with passive observation.

During interaction with virtual objects, the variation in height of adjacent local areas of the virtual contact surface could be assessed in no other way than through kinesthetic and proprioceptive feedback (Coello, 2005) when the hand can move forward or backward. In this case, the body of an observer acts as a frame of reference. When the person moves a stylus across a flat surface of the screen or the touch tablet, the absence

of corresponding hand movements regarding the surface worsens perception of depth cues and representation of 3D objects degrades. On the other hand, when haptic sense cannot provide a high accuracy of alignments concerning pictorial cues such as foreshortening, the person has to exclusively rely on visual cues (Withana et al. 2010). To change the cursor position in the virtual space (along Z-axis) Lapedes with colleagues (2006) proposed to move the tablet PC. However, such an embodiment also changes physiological and pictorial cues, distracts and leads to hand fatigue.

As soon as advanced 3D display technologies will become ubiquitous, they will require new ways of interacting with virtual objects. In the context of touch, the surface limits interactions in two dimensions, limits immersion into a virtual scene. How to extend surface computing interaction beyond the surface? There have been numerous attempts to simulate a physical contact with virtual objects relying on grounded and non-grounded forces. The latter can also constrain users' activity with respect to their own body, the computer screen and input devices (Withana et al. 2010; Kamuro et al., 2009; Kyung and Lee, 2008; Nakadaira et al., 2008; Fiorentino et al., 2005). The non-grounded kinaesthetic device (StickGrip) was developed to provide a true sense of displacements of the stylus grasp area (the grip) regarding the stylus tip. Distance and direction of the grip displacement were coordinated with pictorial cues of the regions of inspection within the image (Evreinov et al. 2009). The goal of the present research was to evaluate of the new stylus-based interaction technique in the situation when visual inspection of images can mislead an observer.

## 2. METHODOLOGY AND RESULTS

### 2.1 Test Images

Partial occlusion, interposition and lightness of the neighboring image areas, surface orientation and the volumetric shapes (concave, convex) might compete with each other (Schlerf et al. 2004, Dresch et al. 2002). The depth cues selected for testing were curvature degree and curvature direction of the simulated shapes, which could resemble a salad bowl (Figure 1). A virtual point of observation defined the imagined spatial position and orientation of the shape.

Five types of images resembling different objects have been simulated using only two overlapping ellipses. The test objects were as follows: convex, concave, inverted convex, inverted concave and displaced flat ellipses (Figure 1). The background of the test images had a luminance level of 128 (of 256 gray levels) and had a size of 330 by 390 pixels. The major external ellipse had been created as 8-bit grayscale GIF image having a size of 300 by 150 pixels and the predefined brightness gradient.

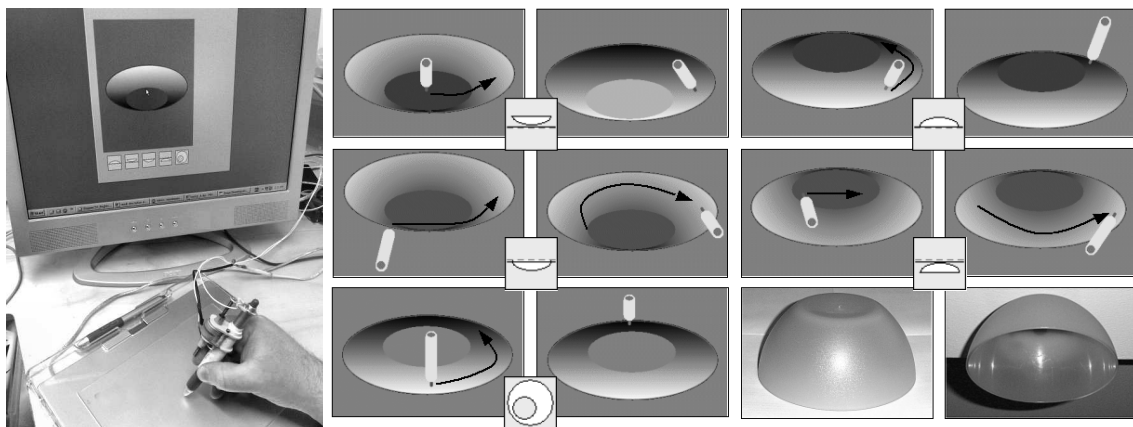


Figure 1. Exploration of images enhanced with a pencil-like visual pointer: black arrows indicate the direction of anticipated exploratory movements, real-world physical models of inverted images and icons to be selected.

Sixteen pictures of major ellipses with different brightness gradients have been stored in a local directory of the testing application and loaded automatically. Four types of achromatic gradients have been used to fill

in the ellipse: linear top-down, linear bottom-up (inverse), sunburst and sunburst inverted. The following foreground-background ratios of the luminance profile were used: 1:16 (or 16:256), 1:8 (or 32:256), 1:4 (or 64:256), 1:2 (or 128:256), where the maximum value (256) corresponded to the white color, and a minimum (1) corresponded to the black color.

The minor ellipse had a predefined size of 150 by 70 pixels, and its color varied from 60 to 120 with a step of 20 or from 180 to 240 of the gray scale level for convex and concave forms. In the case of simulating flat objects, the color of the minor ellipse was set to 128. The color of this ellipse varied from 80 to 160 or from 160 to 220 for inverted shapes. An external lighting condition was stabilized during the test.

Thus, displacement of the minor ellipse in relation to the major ellipse and different brightness gradients of the ellipses should create an illusion of 3D objects. At that, perceptual information delivered by the pencil-like visual pointer and the StickGrip device has been coordinated with pictorial cues.

## 2.2 Apparatus

The StickGrip device is a pen-shaped kinesthetic display that relies on sensations in the fingers such as self-perception of the finger joint-angle positions (Tan et al. 2007). Figure 1 (on the left) illustrates the experimental setup. The lightweight prototype of the motorized pen holder (grip) has been developed which provided physical displacements of the grasp area regarding the tip of the Wacom pen. The StickGrip has a dynamic range of 40 mm ( $\pm 20$  mm) of the grip displacement with an accuracy of ( $\pm 0.8$  mm). An average speed of the grip was about 15 mm/s for the pen having a length of 140 mm. Displacements of the grasp area in this range give a true feedback about the distance and direction (closer and further) regarding the surface of the tablet. Distance and direction of the grip displacements were coordinated with the average brightness of the local area of  $5 \times 5$  pixels of the pointed region and specific orientation of the simulated surface.

## 2.3 Participants

Sixteen volunteer subjects participated in the study (12 males and 4 females). They had normal or corrected-to-normal vision and none of them reported sensitive dysfunction in fingers. The participants were right-handed regular computer users and during the tests they used their right hand to explore the images haptically by moving the pen over the tablet. The age of the participants ranged from 24 to 35 years, with a mean age of 30 years. None of the subjects were aware of the setup before the experiment.

## 2.4 Procedure

The participants of the experiment had to imagine the shape, the point of an observation and, by clicking on the appropriate icon, answer the question: would it be possible to put the shape on the table or to fit it to the ceiling like a ceiling lamp diffuser in a suitable position? The participants were seated at a distance of about 40-60 cm away from the monitor displaying the images of the test. The experiment consisted of three sessions, and the participants were instructed about the procedure of the experiment, which took under one hour to complete. Then an informed consent from each participant was obtained. One trial was given to the subject to become familiar with the sequence of actions required. Each session consisted of 80 trials.

During the first session, the participants made a decision (baseline assumption) about the type of the image presented (convex, concave, inverted convex, inverted concave or displaced flat object) by relying exclusively on passive observation, using a limited number of the pictorial depth cues such as overlapping and brightness gradient. Then, during the second and the third session, an indirect exploration of the images was performed using a tablet pen. The supplementary cues during the second and third session were selected randomly. For instance, during the second or the third session, an exploration of the simulated objects was enhanced with a visual pointer that has appeared as an image of the pencil tangent to the virtual surface of the object. The participants were asked to make inspection visually and interact with objects indirectly through the tablet pen. Thus, they could coordinate the cursor movements over inspected virtual surfaces to collect supplementary information about the curvature gradient and direction indicated by the pencil-like pointer. By adding supplementary visual cues, we could not exclude kinesthetic and proprioceptive feedback signals accompanying the movements of the hand holding the pen. In such a condition according to Faconti et al.



(2000), haptic information could compete with the mental model by worsening or not the visual-kinesthetic imagery related to the image depth and visual perception of curvature.

The inclination angle and direction of the visual pointer was calculated by taking into account the cursor position, the type of the form presented, and a virtual point of observation. A remote point of observation could be located either above or below the virtual object. The length of visual pointer was shortened or lengthened as the function of the average brightness of the local area of 5 by 5 pixels around the pointer:

$$\text{LengthPen} = \text{LengthPen}_{\min} + (\text{aveBrightness} / 5), \quad (1)$$

Where the index “5” and the minimum value (20 pixels) of the length have been found empirically.

During the third or second session, the participants operated by the StickGrip device. The participants indicated their decision by clicking the corresponding icon on the monitor using a tablet. Simulated “displaced flat” objects were equally presented with other images. In each session, the order of the appearance of the simulated images was randomized. There was no time limit for exploration of the ambiguous images and decision-making. To decrease perceptual learning and knowledge transfer, the participants performed three experimental sessions at once. After each session, participants were interviewed about their experiences, problems and strategies.

## 2.5 Results and Discussion

In total, the results on 3840 trials have been collected during an exploration of 5 classes of simulated objects (convex, concave, inverted convex, inverted concave and displaced flat images) with 4 types of achromatic color gradients (linear top-down, linear bottom-up, sunburst and sunburst inverted) and 4 luminance profiles (foreground-background ratios) presented only once during each of 3 sessions (the passive observation, the use of the tablet pen added with a visual pointer and the use of the StickGrip device) by 16 volunteer subjects. Data of observation and exploration of ambiguous images were averaged over all participants.

### 2.5.1 Passive Observation and Interpretation of Ambiguous Images

In the first session, the interpretation of ambiguous images and decision-making were based on passive observation. As reported by the subjects, an exploration of the convex, concave and inverted concave images required a similar effort and time to make an assumption and click the icon. However, the exploration time and decision-making varied from 1.5 to 4.6 seconds with an average of 2.7 s (SD=1.1 s) for concave, 2.9 s (SD=1.0 s) for convex, and 3.0 s (SD=0.9 s) for inverted concave images.

An exploration of the inverted convex images resembling the salad bowl fitted to the ceiling or the ceiling lamp diffuser, had demanded more time and perceptual efforts to distinguish them from concave images. The time required to assess the inverted convex images varied from 1.6 to 3.9 s with a mean of 3.2 s (SD=0.8 s). An interpretation of the displaced flat images took from 1.5 to 3.8 seconds with a mean of 3.6 s (SD=2.0 s). The participants reported that the displacement of the minor ellipse had a strong influence on their decision.

The investigation of the inconsistency recorded between the expected decision and actual subjective interpretation demonstrated the discrepancy rate from 0 to 16 instances for each of five simulated forms with a mean of 5.5 (SD=4.2). The maximum of discrepancy of 16 instances was recorded during inspection of the inverted concave images. The resulting confusion matrix is shown in Table 1, where the data were averaged over all participants.

Table 1. Confusion matrix for the expected decision and subjective interpretation of the ambiguous images

		Ambiguous images classified as, %				
		Convex	Concave	Inv.convex	Inv.concave	Dsp.flat
Expected decision	Convex	72.47	0	2	20.67	4.86
	Concave	0	71.33	24.67	0	4
	Inv.convex	0	18.67	79.2	0	2.13
	Inv.concave	41.47	0.4	6.2	49.8	2.13
	Dsp.flat	3.4	0.47	0	4.46	91.67

## 2.5.2 Exploration of Images by the Tablet Pen enhanced with Visual Pointer

When the participants made a decision based on the observation and indirect exploration of the images using the tablet pen, the cursor movements were enhanced with the pencil-like visual pointer. The subjects had to rely on pictorial depth cues and actively use supplementary information, such as the inclination angle, the length and direction of the pointer. However, the pointer behavior could not always be consistent with ambiguous pictorial cues. Based on the comments given by the subjects, we were informed that when changes of the pencil-like visual pointer were foreseen and supplementary visual information was matched to pictorial cues, it was well accepted by the observers. Sometimes, from the viewpoint of the subjects, pointer behavior differed from an expected (visual-kinesthetic imagery). Therefore, in ambiguous situations, the subjects often ignored the supplementary information.

The interpretation of all the images observed and explored by the tablet pen, took from 2.7 to 17.2 s with a mean of 7.7 s (SD=4.3 s). The minimum time was required to make a decision in relation to the displaced flat images and the maximum time was required to analyze inverted concave images. An exploration and decision making in relation to the convex images required the time from 3.1 to 13.5 s with an average of about 6.8 s (SD= 3.8 s). The convex images were often confused with inverted concave images, while the position of the pointer's tip was different in these conditions.

The interpretation of the concave images and making response required from 2.2 to 14.3 s with a mean of 7.8 s (SD=4.1 s). The concave images were often mixed up with the inverted convex forms. Some participants commented that when a minor ellipse had a high brightness they were imagining "a lamp diffuser" and stopping thorough exploration of the picture. Due to knowledge transfer, some subjects anticipated watching a dark bottom of the salad bowl, as it was already observed during previous session. When a minor ellipse had a high brightness, the participants were referring to inverted convex images.

An exploration of the inverted convex images took from 3.8 to 11.3 s with a mean of 7.9 s (SD=2.6 s). When the subjects ignored supplementary cues, the inverted convex objects have been confused and associated with the icon designating a concave form (the image of a salad bowl).

The time of investigation of the inverted concave images varied from 3.1 to 17.2 s with a mean of 9.2 s (SD=5.2 s). Similarly, inverted concave images were mixed up when the pictorial cues dominated over supplementary information. During an exploration of the displaced flat images, the time spent by the participants to analyze and interpret visual objects varied from 2.7 to 12.4 s with a mean of 6.8 s (SD=4.3 s). However, the most of faults in such a case were identified as "accidental selections of the wrong icon".

An exploration of images with the tablet pen revealed the minimum of the discrepancy of 0.0 instances between the expected decision and actual subjective interpretation of the ambiguous visual objects, and the maximum of discrepancy of 16 instances, with a mean of 3.6 (SD=4.0) instances. It was demonstrated that during an exploration of inverted concave images the participants made a maximum number of misinterpretations – 13, similarly as during the passive observation. The confusion matrix is shown in Table 2 where the data were averaged over all participants.

Table 2. Confusion matrix for the expected decision and subjective interpretation based on exploration of images with the tablet pen enhanced with visual pointer

		Ambiguous images classified as, %				
		Convex	Concave	Inv.conve x	Inv.concave	Dsp.flat
Expected decision	Convex	81.13	0.47	0	16	2.4
	Concave	0.87	89.33	9.53	0	0.27
	Inv.convex	0	19.33	79	0.87	0.8
	Inv.concave	34.87	8.13	5.87	49.8	1.33
	Dsp.flat	1.8	1.33	0.87	4	92

## 2.5.3 Exploration of Ambiguous Images by the StickGrip Device

The participants also investigated the same five types of ambiguous images with enhanced kinesthetic and proprioceptive feedback provided by the StickGrip device. An exploration and interpretation of images required from 4.4 to 16.8 s with a mean of 10.6 s (SD=3.5 s). The time spent by the subjects to analyze the convex images by the StickGrip varied from 6.6 to 16.7 s with a mean of 10.9 s (SD=4.3 s). Inspection and interpretation of the concave images took from 5.5 to 13.1 s per trial with a mean of 10.4 s (SD=2.8 s).

Decision-making in relation to the inverted convex images required from 7.3 to 13.6 s with a mean of 10.4 s (SD=2.5 s). An exploration of the inverted concave forms varied from 8.8 to 16.8 s with a mean of 12.1 s (SD=3.0 s). To explore and classify the displaced flat images, the participants spent from 4.4 to 15.8 s, with a mean of 9.2 s (SD=4.9 s). Inconsistency between expected decision and actual subjective interpretation of the ambiguous objects varied from 1 to 9 instances with a mean of 2.8 (SD=3.8) instances. The confusion matrix is shown in Table 3, where the data were averaged over all participants.

Table 3. Confusion matrix for the expected decision and subjective interpretation based on observation and exploration of images with the StickGrip device

		Ambiguous images classified as, %				
		Convex	Concave	Inv.convex	Inv.concave	Dsp.flat
Expected decision	Convex	92.33	1	1	4.67	1
	Concave	0	96.67	1.8	0	1.53
	Inv.convex	0.53	15.6	81.87	0	2
	Inv.concave	17.47	2.53	6.67	71.8	1.53
	Dsp.flat	2.47	0	0	3.16	94.37

### 2.5.4 Visual Models, Subjective Interpretation and Discrepancy Analysis

The analysis of observation and exploration of the visual models showed that both the minimum and the maximum number of confusions were revealed for the inverted concave images. Whereas during inspection of images with the additional feedback components, the confusions recorded were less frequent (Figure 2).

The paired-samples t-test indicated a significant difference between the confusion indexes (summarized instances of discrepancy) recorded during passive observation and an exploration of five classes of visual models by the StickGrip device  $t(15) = 3.1719$  ( $p < 0.05$ ), while correlation was high and significant 0.94 ( $p < 0.05$ ). The paired-samples t-test also showed that there was a significant difference between the confusion indexes recorded during an exploration of images by the StickGrip device and the tablet pen  $t(15) = 3.0719$  ( $p < 0.05$ ), but correlation of these two variables was less and not significant 0.71 ( $p > 0.1$ ). No significant difference was found between the passive observation and inspection of images by the tablet pen enhanced with the visual pointer  $t(15) = 1.908$  ( $p > 0.1$ ). Though correlation of the confusion indexes under these conditions was about 0.88, it was not significant ( $p > 0.05$ ).

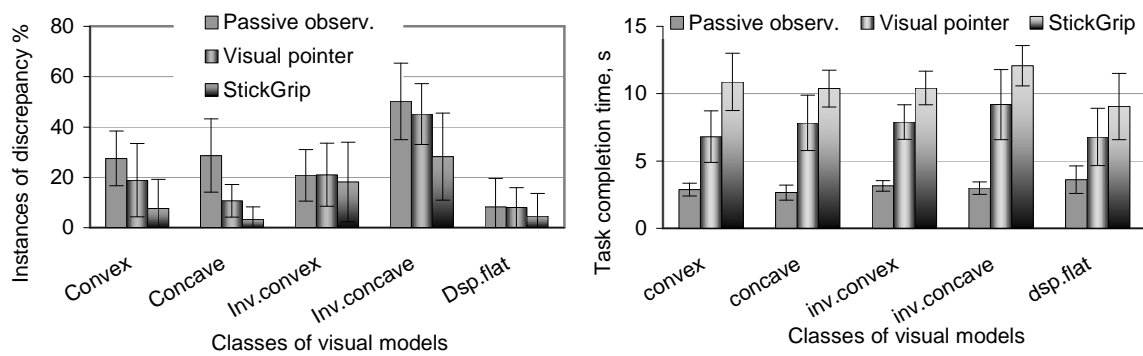


Figure 2. The summarized instances of discrepancy between the expected decision making and actual subjective interpretation of five classes of ambiguous visual models and the mean time under three experimental conditions

The task completion times the participants spent to evaluate the ambiguous visual models under three experimental conditions, are shown in Figure 2 (on the right). The paired-samples t-test within the data presented, revealed that there was a statistically significant difference between the task completion time during passive observation and an exploration of visual models performed by the tablet pen enhanced with the visual pointer  $t(15) = -8.9539$  ( $p < 0.001$ ), during passive observation and an inspection of visual models performed by the StickGrip device  $t(15) = -12.5605$  ( $p < 0.001$ ), during visual inspection of visual models by the tablet pen enhanced with the visual pointer and by the StickGrip device  $t(15) = -8.9873$  ( $p < 0.001$ ). There was no significant correlation detected between these variables, the correlations were as follows: -0.34 ( $p > 0.5$ ), -0.60 ( $p > 0.1$ ) and 0.77 ( $p > 0.1$ ) accordingly.

The results demonstrated that when the participants made visual interpretation of ambiguous visual models based on passive observation, they performed faster (3.1 s, SD=1.2 s) while the results (subjective interpretation of the images) were often confused and different from the expected. An average discrepancy summarized over five classes of visual models was of about 27.1%. An inspection of visual models with additional pictorial cues (the pencil-like visual pointer) required spending more than twice the amount of time to explore an image and make a decision (7.6 s, SD=4.3 s). The summarized average discrepancy between the expected and actual interpretation of the images decreased to 21.75% depending on their ambiguity.

Because a precise stepper motor has limited the transition time of the grip, an inspection of the images with additional haptic components took more time of about 7.8 s (SD=4.1 s), while an average discrepancy between the expected and actual interpretation of the images summarized over five classes of visual models decreased to 12.6%. The participants could spend more time during passive observation to be more accurate in making their decisions. However, when the long thinking took place with a mean of 9.6 s (SD=5 s) during the passive observation of visual models, the number of discrepancies between the expected and actual interpretation of ambiguous images increased up to 52.4% (Figure 3). The reason for that is that the visual inspection time could also influence the judgments of an observer. According to Kjelldahl (2003), short-time visual inspection of images could reinforce perception of discrete cues, and make the subject more sensitive to additional feedback components.

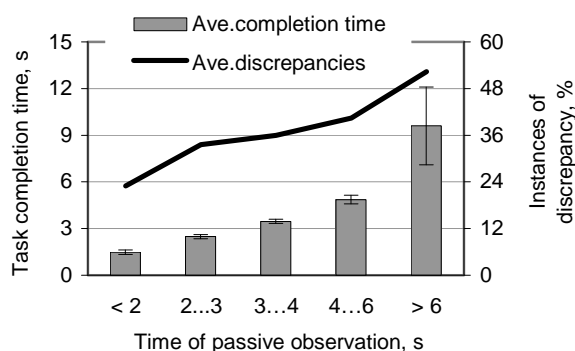


Figure 3. Conformity between the task completion time and decision-making during passive observation of ambiguous visual models, index of correlation 0.957 ( $p < 0.01$ )

### 3. CONCLUSIONS

In the case of interacting with ambiguous images, pictorial information needed to be clarified. The goal of present research was to examine whether the kinesthetic and proprioceptive sense of displacements of the stylus holder (forward or backward) with respect to the stylus tip could improve visual information and understanding of ambiguous images. The ambiguous visual models were achromatic images composed from only two overlapping ellipses with different brightness gradient and relative position presented on a flat screen. These visual models had to simulate convex, concave and displaced flat images resembling realistic objects in a specific orientation. The ambiguous images were evaluated through passive observation and relying on additional feedback components. The supplementary information was delivered by the tablet pen enhanced with the pencil-like visual pointer and by the StickGrip device.

When inspection of images was carried out using the tablet pen enhanced with the pencil-like visual pointer, the subjects had to rely on pictorial depth cues and actively use supplementary information such as the inclination angle, the length and direction of the pointer. An exploration of images with the visual pointer required more than twice the amount of time to complete the task. Distinguishability of images in such a case decreased discrepancy between the expected and actual decision-making only by about 2.6 for concave and by about 1.3 for convex images.

During interpretation of the convex images ambiguous with their inverted concave counterparts, an exploration carried out by the StickGrip device showed threefold better result than with the use of the pen

enhanced with the pencil-like visual pointer. Interpretation of the concave images versus inverted convex counterparts was five times more accurate with the use of the StickGrip device.

The StickGrip device is considered as a robust tool having the potential for everyday work with graphic editors, e.g., to engineers, architects, interior designers and ordinary users. In further research, we plan to examine pattern-based multi-point inspection and interaction with virtual surfaces.

## ACKNOWLEDGEMENT

This work was supported by the Academy of Finland (grant 127774).

## REFERENCES

- Boyd, D., 2001. Depth Cues in Virtual Reality and Real World: Understanding Differences in Depth Perception by Studying Shape-from-shading and Motion Parallax. Undergraduate computer science honors thesis at Brown University, USA, CA.
- Coello, Y., 2005. Spatial context and visual perception for action. *Psicológica*, Vol. 26, pp. 39-59.
- Dresp, B. et al, 2002. Depth perception from pairs of overlapping cues in pictorial displays. *Spatial Vision*, Vol. 15 No. 3, pp. 255-276.
- Ernst, M.O. et al, 2002. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, Vol. 415, pp. 429-433.
- Evreinov, G. et al, 2009. Method, computer program and device for interacting with a computer. Finland Patent app. G06F ID20090434.
- Faconti, G. et al, 2000. Haptic Cues for Image Disambiguation. *Computer Graphics Forum*, Vol. 19, No. 3, pp. 169-178.
- Fiorentino, M. et al, 2005. The Senstylus: a novel rumble-feedback pen device for CAD application in Virtual Reality. Proceedings of WSCG, Plzen, Czech Republic, pp. 131-138.
- Flick, P., 1962. Ten tests of the visual-haptic aptitude. *Studies in Art Education*, Vol. 4, No. 2, pp. 24-34.
- Grossberg, S., 1994. 3D vision and figure-ground separation by visual cortex. *Percept. and Psychophysics*, Vol. 55, pp. 48-120.
- Grossberg, S., 1997. Cortical dynamics of three-dimensional figure-ground perception of two-dimensional pictures. *Psychological Review*, Vol. 104, 618-658.
- Kamuro, S. et al, 2009. Pen de Touch. Proceedings of 36th Int. Conf. on Computer Graphics and Interactive Techniques. ACM SIGGRAPH. Emerging Technologies, New Orleans, USA. Poster, Article 51.
- Kjelldahl, L., 2003. A survey of some perceptual features for computer graphics and visualization. Proceedings of Annual SIGRAD Conference. Umeå University, Umeå, Sweden, pp. 5-10.
- Kyung, K. and Lee, J., 2008. wUbi-Pen: windows graphical user interface interacting with haptic feedback stylus, ACM SIGGRAPH, New Tech Demos. Los Angeles, California, USA.
- Lapides, P. et al, 2006. The 3D Tractus: A Three-Dimensional Drawing Board. Proceedings of the First IEEE Int. Workshop on Horizontal Interactive Human-Computer Systems. Adelaide, South Australia, pp. 169-176.
- Mensvoort van, K., 2009. What you see is what you feel. PhD Thesis. Eindhoven University of Technology Library, Netherlands.
- Nakadaira, A. et al, 2008. 3D Pointing Method, 3D Pointing Device, 3D Display Control Device and 3D Display Control Program. US Patent app. 2008/0225007. G06F 3/038.
- Schlerf, J. et al, 2004. 3D shape-contingent processing of luminance gratings, *Vision Research*, Vol. 44, pp. 1079-1091.
- Tan H.Z., et al, 2007. Discrimination and Identification of Finger Joint-Angle Position Using Active Motion. *ACM Transactions on Applied Perception*, Vol. 4, No. 2, Article 10.
- Wijntjes, M.W.A., et al, 2009. Haptic perception disambiguates visual perception of 3D shape. *Experimental Brain Research*, Vol. 193, No. 4, pp. 639-644.
- Withana, A., et al, 2010. ImpAct: Immersive haptic stylus to enable direct touch and manipulation for surface computing. *ACM J. Computers in Entertainment*, Vol. 8, No. 2, Article 9.

# A SYSTEMATIC EVALUATION OF MOBILE SPREADSHEET APPS

Derek Flood<sup>1</sup>, Rachel Harrison<sup>1</sup>, Clare Martin<sup>1</sup> and Kevin McDaid<sup>2</sup>

<sup>1</sup>*Oxford Brookes University*

<sup>2</sup>*Dundalk Institute of Technology*

## ABSTRACT

The power and flexibility of spreadsheets have made them an essential part of modern business. The increasingly mobile nature of business has created a need to access spreadsheets while on the move. Mobile devices such as the Apple iPhone and Blackberry have enabled users to do this but the small nature of these devices has caused a number of issues for mobile spreadsheet users. This paper presents a systematic evaluation of mobile spreadsheet apps available on the iOS platform which not only includes an examination of the range of available features and functions but also examines the usability of these applications. This work also recommends some ways in which the usability of mobile spreadsheet apps can be improved.

## KEYWORDS

Mobile Apps, Usability, Mobile spreadsheets, spreadsheets

## 1. INTRODUCTION

The spreadsheet application has emerged as one of the most important tools within the financial sector. Research (Croll, 2005) has shown that spreadsheets are an integral part of the day-to-day operations of many businesses within this sector. Consequently there is a growing need for access to the information contained within spreadsheets while on the move.

The power of mobile devices is continually increasing and users are able to accomplish more tasks using these devices. Mobile devices can now enable users to access spreadsheets while on the move, however the limitations of these devices, such as small screen size, limited connectivity and limited processing power, have caused a number of usability issues.

Section 2 of this report highlights some of these issues. Modern spreadsheets contain a large number of cells and often can spread across multiple worksheets, making it difficult for users to navigate the spreadsheet. Looking at such spreadsheets on a mobile device further exacerbates this problem because of the small screen size required by mobile devices for portability.

In order to determine which of the available spreadsheet apps available on the iOS platform is most suitable for viewing such spreadsheets, a strict protocol was developed for evaluating these apps in terms of both functionality and usability. Section 3 of this report outlines this protocol and details how suitable spreadsheet apps were identified.

Section 4 presents the results of the evaluation and highlights some of the common issues among these apps. By identifying issues in this way it is believed that solutions can be found that would further enable spreadsheet app developers to provide more usable apps. Section 5 presents a number of recommendations that have arisen out of the evaluation of these apps and in particular the common issues that exist. Section 6 then outlines the limitations of this work and how these limitations will be addressed in the future. Section 7 then concludes this paper.

## 2. RELATED WORK

The rapid progression of technology has led to an increase in the number of mobile applications available. Although these applications offer a number of advantages in terms of portability and convenience they do so at the cost of usability. Zhang and Adipat (2005) have highlighted a number of issues that affect the usability of mobile applications:

- **Mobile Context:** When considering mobile applications the user is not tied to a single location. This will also include interaction with nearby people, objects and environmental elements which may distract a users' attention.

- **Connectivity:** With mobile devices connectivity is often slow and unreliable and therefore will impact the performance of mobile applications which utilise these features.

- **Small Screen Size:** In order to provide portability mobile devices contain very limited screen size meaning that the amount of information that can be displayed is drastically reduced.

- **Different Display Resolution:** The resolution of mobile devices is reduced from that of desktop computers resulting in a lower quality images.

- **Limited Processing Capability and Power:** In order to provide portability, mobile devices often contain less processing capability and power. This has the effect of limiting the types of applications that are suitable for mobile devices.

- **Data Entry Methods:** The input methods available for mobile devices are difficult and require a certain level of proficiency. This problem increases the likelihood of erroneous input and decreases the rate of data entry.

As spreadsheets are used in a variety of ways, they can contain vast amounts of data as well as tables, charts and graphs. These spreadsheets can become quite large which can be difficult to navigate even on a conventional desktop computer. This usability problem is exacerbated on mobile devices by their limitations. The small screen size requires the user to perform considerably more navigation when looking at large spreadsheets. This may cause users to find it difficult to conceptualise the overall spreadsheet and to see how the section on-screen fits with this overall picture.

Flood et al. (2008), have identified navigation as an issue that affects the performance of people debugging spreadsheets through voice recognition technology. By addressing this issue it was found that the performance of users debugging spreadsheets could be increased. It was also found that participants audited more cells with the improved navigation system, which is an important aspect of the spreadsheet debugging process.

Mobile devices generally do not contain a traditional keyboard as the size required would be too large to enable portability. Some devices incorporate a physical keyboard which utilises small keys while other devices use touch screen technology to present a keyboard to the user on screen. These keyboards also require the physical keys to be smaller than traditional keyboards to fit all keys on screen.

Chen et al. (2010) conducted an evaluation of users entering text on a small size QWERTY keyboard. This evaluation required 15 participants to enter a passage of text using a small sized keyboard. On average participants used 540 keystrokes to enter the passage of text. The most prevalent type of error made by these participants during the task was a key ambiguity error, which occurred when a user entered a character other than the target character. It was found that on average, participants made 9.33 key errors on the first typing task. It is also worth noting that all participants made at least one error of this type during the study.

Errors of this type, when made on a spreadsheet, may result in a misspelled word or in an incorrect reference in a cell formula, which could alter the bottom line value of a spreadsheet substantially. It has been shown repeatedly that even on desktop computers errors like this persist. Powell, Baker and Lawson (2009) have found that over 85% of operational spreadsheets contain errors. This study furthers previous research by Panko (1998) which found similar levels of spreadsheet error.

The limited processing power of portable devices has meant that existing spreadsheet applications may not function correctly when run on these devices. In an attempt to address this issue, developers have created a number of spreadsheet apps which scale down the level of functionality to enable users to view and use spreadsheets in a mobile context. Most of these applications however, are limited in terms of functions available and spreadsheet size.

### 3. EVALUATION PROTOCOL

The iOS operating system runs on most portable Apple devices such as the iPhone, the iPad and the iPod touch. It was decided to focus on this platform initially as Apple are one of the leading providers of mobile devices. It is planned to extend this evaluation to other platforms such the Blackberry OS and Google Android. Apps for the iOS platform are distributed through the app store maintained by Apple, and can be downloaded directly on to the handset. In order to evaluate these apps a strict evaluation protocol was developed and is presented below.

1. *Identify all potentially relevant applications.* There are a number of ways to conduct a search for appropriate applications, including a standard web search, and current software distribution methods make this increasingly easy. Most of the major mobile phone platforms now have an associated online application store. As this work is focused on spreadsheet apps for the iOS operating system, a search of the App store was conducted. The search string “Spreadsheet” was used during the search.

2. *Remove light or old versions of each application.* Many software developers release trial versions of their systems, which are often free. Some of these versions include only a subset of the functionality offered by the full application whilst others allow full access to the application but for a limited time period. These types of applications should be removed if the full version of the app is also included within the search results.

3. *Identify the primary operating functions and exclude all applications that do not offer the required functionality.* The primary operating functions include frequently used functions and also occasionally used functions that are essential for the correct operation of the system in a desired context. For example, the initial system setup might include language and currency settings that would depend upon the country of use. The primary functionality of interest is to allow a user to perform spreadsheet tasks on a mobile device. This does not include the ability to access online spreadsheets as an internet connection cannot always be guaranteed.

4. *Identify all secondary functionality within the remaining apps.* In addition to the primary operating functions, mobile apps will offer users a range of secondary functionalities which can enhance the application. A thorough knowledge of these functions will enable the application developers to see what functionality is available and may present opportunities for additional functionality to be included in future applications.

5. *Install each of the remaining applications, and test each of the tasks using*

a. *Keystroke level modelling.* Keystroke Level Modelling (KLM) is a well established technique (Card et al., 1983) for estimating the time taken to complete certain tasks. This will provide a quantitative measure of the difference between the efficiency of applications. Time estimations are based on the number of interactions required to complete the task. The time for each interaction may vary between individuals. KLM was used to measure the average number of interactions required to enter each of a set of 5 single digit numbers as well as the number of interactions to enter a subsequent formula to total these five numbers.

b. *Heuristic evaluation.* Standard heuristics (Nielsen and Molich, 1990) can be used to evaluate desktop applications. These have been modified for the evaluation of mobile applications (Bertini et al., 2008) so that contextual information can be taken into account.

## 4. RESULTS

### 4.1 App Search

After searching the App store 23 potential spreadsheet apps were identified. After removing light and free versions of other apps and those that did not allow users to create or view spreadsheets offline 12 spreadsheet apps remained. A full list of the apps evaluated can be seen at <http://tech.brookes.ac.uk/groups/pacmad/>

The App store allows users of an app to give each app a rating between 0 and 5. Figure 1 shows the number of apps by rating. It can be seen that the majority of spreadsheet apps have been given a rating of 2 or under. This indicates that users, in general, did not rate these apps very highly. It should be noted that 3 of the apps had not received enough ratings for an average rating to be produced and are therefore not included. It



can also be seen that none of the applications were given an average rating of 2.5. The highest rated app was *Spreadsheet* by AppAauthours receiving an average user rating of 4.0.

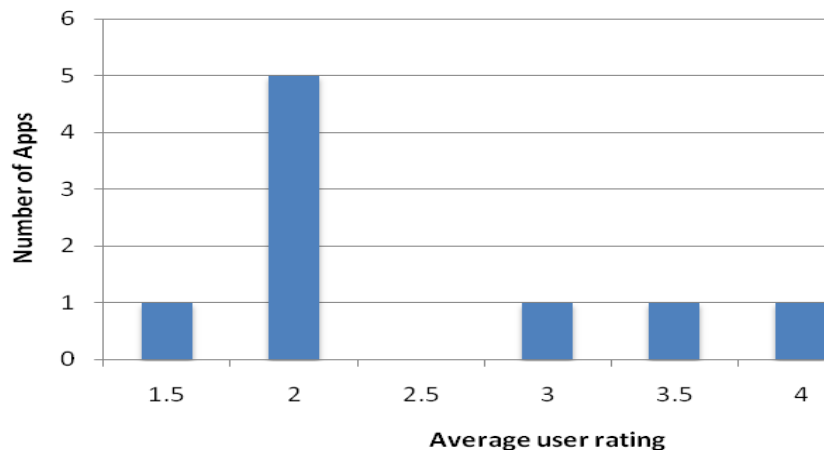


Figure 1. Frequency of App Ratings

## 4.2 App Findings

### 4.2.1 Transferring Data

Table 1. Percentage of apps by methods of transferring data

	Copying to the device	Copying from the device
Email	41.67%	58.33%
Wi-Fi	58.33%	58.33%
Google Integration	25.00%	25.00%
iTunes	58.33%	58.33%

Transferring data is an important aspect of mobile spreadsheet apps. A user may wish to access an existing spreadsheet on a mobile, during a meeting for example or while travelling. If the user subsequently changes this spreadsheet they may wish to then transfer the spreadsheet back to the original location once it has been updated. Mobile spreadsheets primarily offer 4 ways in which spreadsheets can be transferred between a mobile device and a desktop or laptop computer: Email, Wi-Fi, Google Integration and via the iTunes application.

One of the most convenient features of mobile devices is that they provide access to emails while on the move. A common way of transferring digital content, such as spreadsheets, is via email. It would be therefore desirable for mobile spreadsheet apps to allow users to open a spreadsheet on the mobile directly from an email. Approximately 42% (5) of the evaluated apps allow users to do this. If a user has altered the spreadsheet they can save the changes by emailing the completed spreadsheet directly from inside the app. 58.33% (7) apps allow users to do this.

As a number of mobile devices feature Wi-Fi capability, 58.33% (7) mobile spreadsheet apps allow users to transfer files in this manner. Once the mobile device is connected to the same network as the desktop computer, the user can mount the mobile device as an additional drive and copy spreadsheets to and from the device. Some of the spreadsheet apps allow a user to connect to the device through the IP address of the device and access the files through a web based interface.

A quarter of the apps reviewed allowed users to access spreadsheets stored on the Google spreadsheets online account. This facility requires the mobile device to be connected to the internet, which may not always

be available. It also assumes that users have a Google account and their spreadsheets are stored on Google servers. These applications may offer similar facilities to other online storage solutions such as drop box.

The final method of transferring data to and from mobile devices is through the iTunes application. Apple devices, such as those running the iOS, transfer data to and from the device through the iTunes application. Through the 'Apps' tab of this application, 58.33% of the mobile apps evaluated allow users to transfer spreadsheets to and from the device.

The *iTables* application does not allow users to transfer files to the device. It only allows them to create and access spreadsheets on the device.

#### 4.2.2 KLM Analysis

The spreadsheet apps evaluated allow users to create spreadsheets on a mobile device. Due to the limitations of these devices, the way in which users interact with the apps varies. In order to evaluate which app features the most efficient interaction method, the same spreadsheet was created using each of the apps. The number of keystrokes required to create the spreadsheet was recorded for each of the twelve apps and is presented below.

Table 2. Number of keystrokes by task

App	Total Keystrokes	Average Numerical Entry	Entering Sum Formula
Spreadsheet	19	2	4
Sheet <sup>2</sup>	19	2	4
Beepad	19	2.4	4
Docs to go	26	2.2	7
Cellroid G	26	2.2	8
iSpreadsheet	27	2.2	10
Discount Spreadsheet	28	3	9
Quick Sheet	29	2.2	9
Spreadsheet LX	29	3	10
i123	33	3	13
Mariner Calc	34	3.8	10
iTables	36	4	11

As can be seen in Table 2 the *Spreadsheet* app required the fewest keystrokes to create the spreadsheet. This application contained a number of features which reduced the number of interactions required to complete the task. Firstly by default the user is located in cell A1 and can enter data by simply opening the keyboard and entering the desired text. A number of applications, including *Spreadsheet*, open the default iPod keyboard to numerical input as numbers are the most common type of input.

The *Spreadsheet* app also provides quick access to the sum function. As the cursor is placed in cell A6, the user only needs to press the 'fx' button, press 'Sum' and then press 'Done'. The first function in the list is the Sum function and will recognise by default that the user wants to total the numbers located in the cells above the currently selected cell.

The most interactions are required for the *iTables* app. The main reason for the relatively high number of interactions required is how the system treats data input. When the user wants to enter data they first click on the cell they wish to enter the data into. The default character keyboard is presented which will require the user to manually change to the numerical keyboard. In a number of spreadsheet apps the system will allow the user to select another cell while the keyboard is still open but *iTables* does not. It automatically closes the keyboard when the user selects another cell requiring the user to reopen the keyboard and to reselect the numerical keyboard.

The Average numerical entry shows the average number of keystrokes required to enter a number into a cell. As each number was entered to the cell sequentially, it was possible to leave the keyboard on screen during the entry of all the numbers, reducing the number of interactions required for the second and subsequent cells. One or two additional keystrokes are required for the first cell entry depending on the app. The minimum number of interactions required to enter a number is two. The first interaction is to select the cell while the second interaction is to enter the desired digit.

Some apps such as *Sheet<sup>2</sup>* and *Beepad* use the context of the cursor to determine the parameters for functions like the Sum function, enabling the user to enter such a function in three or four interactions. Other apps require the user to manually select these parameters and often require the user to open the keyboard (which needed to be closed to access the functions) to select the ‘:’ character to indicate all cells between the first selected and last selected are included as input to the function. All applications except *i123* and *iTables*, feature a menu which allows the user to see and select all available functions. In the case of both the aforementioned applications, the user is required to manually type in the function name and parameters.

### 4.2.3 Incorporated Functions

A function performs a set of mathematical transformations on a series of inputs to produce an output. Spreadsheet applications contain a large number of pre-implemented functions which allow spreadsheet users an easy way to perform complex calculations. The inclusion or exclusion of these functions can impact both the effectiveness and efficiency of users completing their desired task through the spreadsheet app.

Table 3. Number of Functions by App

Spreadsheet	104	MarnerCalc	145
i123	27	Spreadsheet LX	51
Cellroid G	61	iSpreadsheet	13
Beepad	21	Discount Spreadsheet	23
Documents to go	112	Sheet <sup>2</sup>	146
Quicksheet	80	iTables	1

Table 3 shows the number of functions that are included within each of the spreadsheet apps evaluated. It can be seen that a number of the spreadsheet apps contain only a small number of functions, which will limit the number of pre-existing spreadsheets that can be viewed and manipulated on the mobile device. The *Sheet<sup>2</sup>* features the most extensive range of features incorporating 146 functions, while the *iTables* app only features the sum function.

While evaluating the types of functions available on mobile apps, it was found that the range of specialist functions, such as statistical or financial was quite limited. Ten of the fourteen spreadsheet apps contain financial functions, however most of these apps only feature 5 or less financial functions. It was also found that only two of the spreadsheet apps contain base conversion functions, which enable users to convert between different number bases.

Although there are a lot of similarities between the functions available on mobile apps, it was found that different apps implemented different functions. For example *Sheet<sup>2</sup>* offers users most functions, however it does not include a large number of financial functions that are offered by *MarnerCalc*. However, it should be noted that *MarnerCalc* does not offer any Base Conversion functions.

### 4.2.4 Spreadsheet Features

Spreadsheets are often used for more than just performing complex calculations. A number of features exist within spreadsheet applications that allow users to see beyond the numerical data. Charts and graphs allow users to see different representations of data which can make identifying patterns easier and can also help in allowing users understand what the data means. Only 25% (3) of the spreadsheet apps evaluated featured this functionality. These applications feature only a small subset of the chart types that are found in desktop spreadsheet applications.

In addition to the ability to view chart data, spreadsheet users often use the Sort feature to arrange data in a particular order. This feature is most useful when a large amount of data is stored on the spreadsheet in a table and users need to evaluate the data in a particular order or in groups based on a particular attribute. Of the twelve apps evaluated 50% (or 6 apps) allowed users to sort the data.

The final spreadsheet feature considered was the freeze pane. Freeze panes allow a user to keep a certain number of rows or columns on screen while the rest of the spreadsheet scrolls down or across. This is usually done to keep headings visible while the user is looking through the data. Of the evaluated applications, 33% (4 apps) contained this feature.

One feature that was noticeably absent from all of the apps was Filters. When dealing with large lists of data, filters can enable users to reduce the amount of data that appears on screen, to only rows that meet a

specific criterion. On a mobile device, where the screen space is limited, any feature that would allow users to focus on only the data that is relevant would be advantageous.

## 5. RECOMMENDATIONS

During the evaluation of the spreadsheet Apps, a number of usability issues were observed. The spreadsheet paradigm is just one example of a mobile app however; the issues observed can be seen in other types of mobile app. The following guidelines, if implemented could help to improve the usability of these apps.

- **Consider the big picture.** In some cases the spreadsheet being viewed can be larger than the screen size available and therefore it can be difficult for users to navigate to different sections of the spreadsheet. The same issue has been seen on large web pages, large images and large text documents. In these cases the use of a mini-map can be used to show the user the entire spreadsheet, on a small scale, with the area that is currently displayed on screen highlighted. This approach would give the user an idea of where on the current worksheet they are looking.

- **Consider the type of input.** The most common form of input for spreadsheet cells is numeric or formulaic. A number of the spreadsheet apps however, open the text based keyboard when a user indicates they would like to enter data onto the spreadsheet. If a user does want to enter numeric data into the cell they need to switch keyboards before they can do so. Mobile app developers should consider the type of input they are expecting, textual, numeric or special characters and open the keyboard to this type of entry.

## 6. THREATS TO VALIDITY AND FUTURE WORK

For the evaluation presented here applications were selected for the Apple iOS. Devices that contain this operating system only account for part of the entire mobile application market. The Google Android and Blackberry collection of smart phones also feature a range of mobile spreadsheet applications not available on the iOS. It is therefore intended to extend the evaluation to include spreadsheet apps available for these smart phones.

Part of the protocol presented above requires the use of heuristics to evaluate the usability of the apps. Heuristics of this nature are usually rated by multiple usability experts and these ratings are then compared to identify the most prevalent usability concerns of the application. The validity of this work could be improved by increasing the number of usability experts involved.

During the KLM analysis it was not possible to effectively record swipe actions on the device. This type of action was used on multiple applications to move the visible area of the screen during the creation of the spreadsheet. This was seen consistently across all apps and therefore does not affect the results presented here.

## 7. CONCLUSIONS

This report has evaluated twelve spreadsheet apps available for mobile devices featuring the iOS operating system, such as the iPhone, iPod touch or iPad. These apps provide users with access to spreadsheets while away from a desktop or laptop computer. A number of issues with mobile devices such as small screen size and limited processing power have caused these apps to change the way users interact with them from traditional spreadsheets.

This evaluation has found that there is a wide range of differences between mobile spreadsheet apps. Some apps allow users to view existing spreadsheets on a mobile device while others only allow users to create new spreadsheets in a mobile context. The method by which the user creates the spreadsheet also changes between apps. A KLM evaluation of the apps has shown that the number of keystrokes required to create a simple spreadsheet can vary by as much as almost 90%.

This evaluation has also identified a number of guidelines that app developers should follow to improve the usability of mobile Apps. It has been found that a number of apps do not optimise the input method for

data, therefore complicating the way in which the user enters data. By considering the most common type of data to be entered into a cell, the developers could optimise this.

## REFERENCES

- BERTINI, E., CATARCI, T., DIX, A., GABRIELLI, S., KIMANI, S. & SANTUCCI, G. (2008) Appropriating Heuristic Evaluation Methods for Mobile Computing. *International Journal of Mobile Human Computer Interaction*, 1.
- CARD, S. K., THOMAS, T. P. & NEWALL, A. (1983) *The Psychology of Human-Computer Interaction.*, London: Lawrence Erlbaum Associates.
- CHEN, T., YESILADA, Y. & HARPER, S. (2010) What input errors do you experience? Typing and pointing errors of mobile Web users. *Int. J. Hum.-Comput. Stud.*, 68, 138-157.
- CROLL, G. J. (2005) The Importance and Criticality of Spreadsheets in the City of London. *European Spreadsheet Risks Int. Grp. (EuSprIG)*.
- FLOOD, D., DAID, K. M., CAFFERY, F. M. & BISHOP, B. (2008) Evaluation of an Intelligent Assistive Technology for Voice Navigation of Spreadsheets. *Proc. European Spreadsheet Risks Int. Grp. (EuSprIG)* London, UK.
- NIELSEN, J. & MOLICH, R. (1990) Heuristic evaluation of user interfaces. *the Conference on Human Factors in Computing Systems (CHI90)*,.
- PANKO, R. R. (1998) What we know about spreadsheet errors. *J. End User Comput.*, 10, 15-21.
- POWELL, S. G., BAKER, K. R. & LAWSON, B. (2009) Errors in Operational Spreadsheets *Journal of Organizational and End User Computing (JOEUC)*, 21.
- ZHANG, D. & ADIPAT, B. (2005) Challenges, Methodologies, and Issues in the Usability Testing of Mobile Applications. *International Journal of Human-Computer Interaction*, 18, 293 - 308.

# TOWARDS EVALUATING HUMAN-INSTRUCTABLE SOFTWARE AGENTS

Robert D. Grant<sup>1</sup>, David DeAngelis<sup>1</sup>, Dan Luu<sup>1</sup>, and Dewayne E. Perry<sup>1</sup> and Kathy Ryall<sup>2</sup>

<sup>1</sup>*Empirical Software Engineering Laboratory  
Center for Advanced Research in Software Engineering  
The University of Texas at Austin, Austin, TX 78712*

<sup>2</sup>*Advanced Information Technologies  
BAE Systems, Burlington, MA 01803*

## ABSTRACT

The Bootstrapped Learning (BL) project is an attempt to create software agents (e-students) that are instructable by human teachers through natural instruction methods [Oblinger, 2006]. In this paper, we present an introduction to BL and three years of case studies investigating the use of human subjects in evaluating e-students. In our studies we investigate human teachers' expectations of e-students and important differences between human and software learners, including the greater semantic understanding of humans, the eidetic memory of e-students, and the importance of various study parameters including timing issues and lesson complexity to human performance.

## KEYWORDS

User Study, Evaluation, Instructable Computing, Electronic Student, Machine Learning

## 1. INTRODUCTION

Bootstrapped Learning (BL) is a DARPA program aiming to create software agents that human instructors teach rather than program [Oblinger, 2006]. Creating a domain-independent learning agent (i.e., e-student) can be viewed as providing a more intelligent, natural user interface for underlying machine learning algorithms. Computational agents with this interface could be trained by domain experts who are not necessarily skilled programmers; this is especially valuable for systems that benefit from being "field-trainable", or specializable to a particular need by end users at a faster rate than is usually supported by a traditional software development lifecycle.

From an HCI perspective, we address two important issues: determining which instruction methods are most important for supporting human instruction of e-students, and developing human benchmarks for evaluating an e-student's success at learning. Our group has investigated these issues through a series of exploratory case studies.

In this paper, we begin by providing an overview of the BL research program as context for our evaluation work. We then present our findings from an initial study investigating how human teachers attempt to instruct e-students. Finally, we present two case studies where we explore using human students to generate benchmarks for experimental evaluation of e-students. We omit details of the domain of the final two case studies, as the testing domain must be kept hidden from those creating the e-students until after e-student evaluation.

## 2. RELATED WORK

Bootstrapped learning provides the context for this HCI research. Specifically, we investigate how humans teach and learn and then apply those findings to understand how humans can teach machines in a natural way.

## 2.1 Bootstrapped Learning Overview

The goal of the Bootstrapped Learning program is to build an e-student that can be taught by a human instructor in the same ways that humans instruct one another. As BL provides natural ways for a human to impart knowledge to a software learner, it does not require programming expertise; human instruction of e-students will make it possible to delegate tasks to computers that cannot be easily delegated today and will enable users to rapidly modify deployed systems.

BL differs from other kinds of machine learning (ML) in several ways. Current ML is primarily a modeling tool; it is used to build models when we know something, but not everything, about some target problem. Current ML is a process of discovery, and requires induction over large datasets over which to induce its models. There is no guarantee that target knowledge will be discovered.

BL allows users to impart the target knowledge in a more direct fashion; however, because it involves “natural” ways to impart knowledge, it does not require programming expertise. BL supports conceptual bootstrapping; it leads to meaningful intermediary levels of learned concepts. E-students learn laddered-curricula in which lessons build on previous lessons, whereas in current ML, learning is generally from unstructured data. Like its human counterpart, an e-student assumes all necessary knowledge is possessed by the instructor, and its goal is to learn using the “same” instruction methods used between humans.

Two teams have been working in parallel to explore this new learning paradigm. Our team, the Curriculum Team, is developing BLADE (Bootstrapped Learning Analysis and Curriculum Development Environment). This research includes developing a framework to support BL, a set of laddered curricula across a variety of domains as testing vehicles for the e-student, and an evaluation of the e-student on both hidden and known domains. A separate Learning Team is developing an e-student incorporating several learning strategies [C. Morrison and D. Bryce and I. Fasel and A. Rebguns, 2009].

BLADE includes three agents, whose interactions and relationships are shown in Figure 1. A teacher agent serves as a proxy for an eventual human teacher, instructing and testing the e-student. The student agent is the embodiment of the e-student, typically employing a number of learning algorithms. The world agent serves as a proxy for a domain simulator. Over the first three phases of the BL program the Curriculum Team has developed a set of laddered curricula in a variety of complex domains including Blocksworld [Berland and Perry, 2009], unmanned aerial vehicles (UAV), diagnosis tasks for an international space station (ISS), armored task force maneuvers (ATF), planning robotic arm movements, and a hidden domain.

BLADE uses IL (InterLingua) and ITL (InteracTion Language) [Curtis, 2009], developed specifically for BL, to pass messages between agents in the BLADE framework. For evaluation purposes an automated teacher agent is used to ease scaling and reproducibility. Part of our research is to explore how to best incorporate a human teacher. In a parallel effort we are developing a tool to support human-/e-student instructional interactions, in part informed by the evaluations described in this paper.

We expect the eventual outcome of our research to have impact in two useful ways. First, we will demonstrate that instructable computing is a valid and successful means of providing learning to an e-student. Second, the system we develop will provide the tools for other groups to pursue research in Bootstrapped Learning. These groups will be able to access our framework and supporting materials, including our BL curricula, our associated research papers and documents, and our software for supporting human benchmarking. Access to these resources will allow other researchers to experiment with the development of their own e-students, and explore UI designs and techniques for human instructors to interact with e-students. This living repository will help catalyze future work in BL just as the Irvine Repository [Frank and Asuncion, 2010] drove machine learning in the 1980s, and as DARPA’s MUC competitions [Grishman and Sundheim, 1996] inspired research in corpus-based approaches to natural language processing.

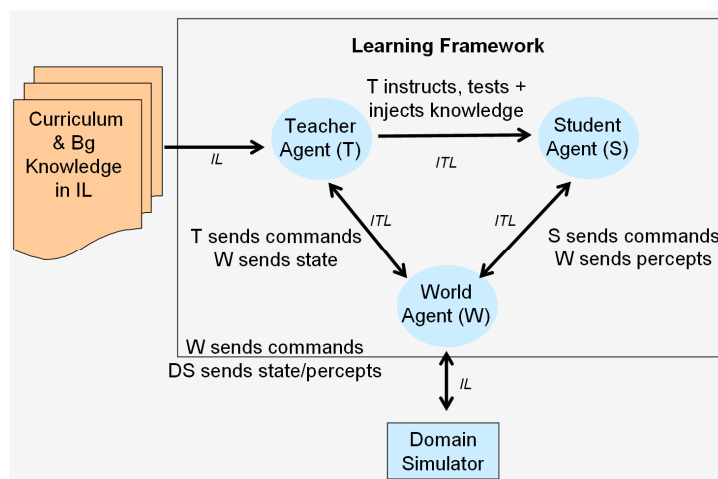


Figure 1. The BLADE Framework

## 2.2 Human Learning and Teaching

Like its human counterpart, an e-student assumes its instructor possesses all relevant capabilities, and its goal is to learn using the same instruction methods used between humans. As part of the Evaluation Team, we are not allowed access to e-student implementation details to ensure our benchmarks are unbiased. We know that the learners are designed to be domain-independent, and that they are specialized to particular Natural Instruction Methods (NIMs) rather than particular problem domains.

The area of computer tutoring can be seen as an inverse problem to what we are investigating. In particular, the area of teachable agents bears some surface similarity to BL. In this field, human students teach learning agents in order to improve their own understanding of concepts (“Learning by Teaching”). One example is the “Betty’s Brain” system [Davis et al., 2003]. However, in these systems the importance is placed on how well the human instructor learns, not on the capabilities of the learning agent.

## 2.3 Human Case Studies

Our human studies use well known techniques from behavioral research, as covered in standard texts such as [Rosenthal and Rosnow, 1991] and [Yin, 2008]. In particular, since little was known about what factors would be critical, our empirical approach is that of exploratory case studies. An exploratory case study is the best approach when little is understood about the subject under study [Yin, 2008]. The intent is to build a deeper understanding of the phenomena in question and to formulate the beginnings of a corresponding theory that can be tested, revised, and expanded with further empirical studies.

## 3. PHASE I STUDY: BLOCKSWORLD

The Curriculum Team performed an initial case study to explore what approaches a human teacher (HT) would take in teaching a particular curriculum to an e-student, what assumptions a teacher would make about a student, and how these assumptions would work in the context of an e-student [Berland and Perry, 2009]. The HTs were asked to consider teaching to the level of a bright two year old. The domain for this study was the Blocksworld simulated environment created by Cycorp, Inc. This environment consisted of a “claw”, or crane-like device which the e-student could control to manipulate “square blocks” and “long blocks”.

In this case study, each of five human teachers first attempted to teach an e-student to construct a stack of three blocks and then to build a simple “doorway” out of blocks in this environment. The target was a simple structure of two stacks of blocks topped with a lintel. We used a Wizard of Oz (WOz) [Dahlback et al., 1993]



style methodology, where the teachers' natural-language instructions were translated into precise terms (IL) for the e-student by a human interpreter.

We compared the methods of the human teachers in terms of type of curriculum taught, teaching time, and how well the e-student performed. We also gathered information from the teachers about their difficulties in the experience and their models of the e-student before and after their teaching session.

All five teachers successfully instructed the student to make a doorway, and more importantly, we gained important insight into how humans attempt to instruct e-students. Some of our observations are listed in Table 1.

Table 1. Phase I Observations

Description	Impact
All HTs ended up using a bottom-up approach to teaching (possibly due to capabilities of the e-student). Some HTs initially used a top-down approach but became frustrated and reverted to a bottom up approach	All BL curricula to date have been authored using a bottom-up structure. In the Phase II and Phase III evaluations with human learners (rather than e-students), we have found that human subjects often prefer a top-down instructional method. It is an open question as to how an e-student might learn with a different lesson structure.
All the HTs overestimated what the e-student knew and could do, assuming knowledge of primitives such as "choose a block" or "look for a clear space."	We believe the domain-independent e-student needs a minimal level of injected knowledge to support "basic human competencies". This can be achieved through the use of background knowledge which can be injected or built into an e-student.
Many of the HTs employed repetition and mnemonics when teaching.	Instructional interfaces should support natural human methods for teaching, including support for informalities. The Curriculum Team is currently exploring this issue.
The HTs differed in their assumptions regarding the linguistic capabilities of the e-student.	Instructional interfaces should mask the linguistic limitations of an e-student and/or a teacher should have a way to query an e-student's capabilities and understanding.

A human teacher teaches differently depending on the target audience [Dahlback et al., 1993]. Phase I was designed to gain insight into how a human teacher would instruct an e-student. Later phases use a standardized curriculum and are more focused on how well the students (human and electronic) learn the curriculum. Incidental benefits from Phase I included a test of the e-student instruction language and a greater understanding of the learning performance of an early version of the e-students.

#### 4. PHASES II AND III: HUMAN COMPARISON STUDIES

In these case studies, the goals were to define and refine requirements, problem solving strategies, and evaluation methodologies for e-students by evaluating a version of the e-student curriculum with human students. We aimed to produce lessons and tests on which human students who scored less than 20% in pre-test could score at least 75-80% in post-test, indicating that learning occurred.

We performed two studies, one in Summer 2009 and one in Summer 2010, which we respectively refer to as "Phase II" and "Phase III". Our overarching goal was to mimic the e-student context as closely as possible in the human studies, as the eventual goal is to directly compare e-student learning with human-student learning on "identical" curricula in controlled experiments.

In all studies, the basic procedure was as follows:

- Introduction and background material presentation
- Pre-test
- Curriculum lessons with web-based quizzes
- Post-test

## 4.1 Natural Instruction Methods

We presented the curriculum to students via three different “Natural Instruction Methods” (NIMs): teaching by telling, teaching by example, and teaching by feedback. Respectively, these consisted of utterances emitted by a teacher, examples performed by a teacher in a domain simulator, and instructions for a student to apply techniques in a simulator with teacher feedback. We abbreviate these lesson types as T, E, and F. In Phase II, we gave each human student all three lesson types, with the option to skip. In Phase III, we tested some students with all three NIMs, and other groups with only one or two NIMs. We also allowed a small group of students to ask questions about the curriculum. We refer to the students who received all three instruction types as the “baseline”.

## 4.2 Evolution of Testing Procedure

As these studies were exploratory, over their course we evolved the details of our testing procedure significantly. At the beginning of Phase II, we began with a direct analog of the relationship between an automated teacher and an e-student. Since we were concerned with evaluating the curriculum and not the subjects, a major concern was preventing a human teacher from unconsciously providing extra-curricular information to the student through facial expressions, gestures, tone, etc.

In this design there was one teacher, one student, and one observer per session. The teacher fed each line of curriculum to the student one-at-a-time through an electronic messaging interface. For the E lessons, a view of the teacher’s domain simulator was duplicated on a monitor visible to the student. For the F lessons, a view of the student’s domain simulator was duplicated on a monitor visible to the teacher, and the teacher provided feedback through the messaging interface. This teaching method was time-consuming, error-prone, and there was no protocol allowing the teacher to report errors or redo instructions. We quickly transitioned to a “self-paced” version of the curriculum.

In this version, instead of a human teacher feeding every line of curriculum to the student and demonstrating simulator usage manually, we formatted the curriculum as PowerPoint slides with instructions and accompanying figures. For the feedback/practice lessons, we provided the student with instructions on procedures to try in the simulator and what the outcome should look like if the procedure was performed correctly; we called this the “choose-your-own-adventure” style. This eliminated complexities in our lab setup, reduced concerns about students learning from extra-curricular cues, and allowed us to run several students in parallel. We began running two subjects at once in Phase II, and with some additional automation we were able to run six subjects at once in Phase III. For more details on our testing procedure see [Grant et al., 2011].

### 4.2.1 Quantitative Results

In these results we focus on a few aspects of our study that are interesting from an HCI perspective. In all results we exclude students that passed the pretest. When we discuss a “post-test score” in the following analysis, we mean the fraction correct out of five randomly-chosen post-test scenarios. When we discuss a realtime post-test, we mean that human students were tested in a domain simulator that automatically advanced through states in realtime. In non-realtime post-tests, the students advanced states manually (though there was still an overall time limit).

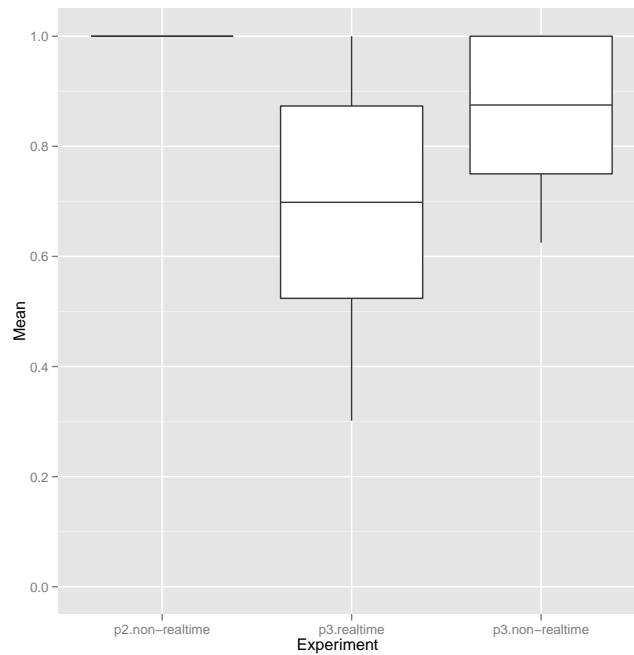


Figure 2. Baseline Plot Test Means

We can compare three groups of baseline subjects: p2.non-realtime, p3.realtime, and p3.non-realtime, denoting students from Phases II and III with and without realtime post-tests, as indicated. These groups contained 28, 12, and 19 subjects respectively. All baseline subjects completed the study in under four hours. Because we increased the difficulty and complexity of the curriculum in Phase III, subjects generally took longer and scored lower than the Phase II subjects (Figure 2). A major revelation in Phase III was the impact of real-time testing on subject post-test scores. We discuss this more in Section 4.2.2.

In Figure 3(a), we show the length of study by NIM-set given. We can see that subjects tended to take less time when given the T NIM, and more time when given the F NIM. In general, the T lessons were shorter than the F lessons, and the F lessons also required subjects to interact with the simulator and encouraged subjects to repeat if necessary.

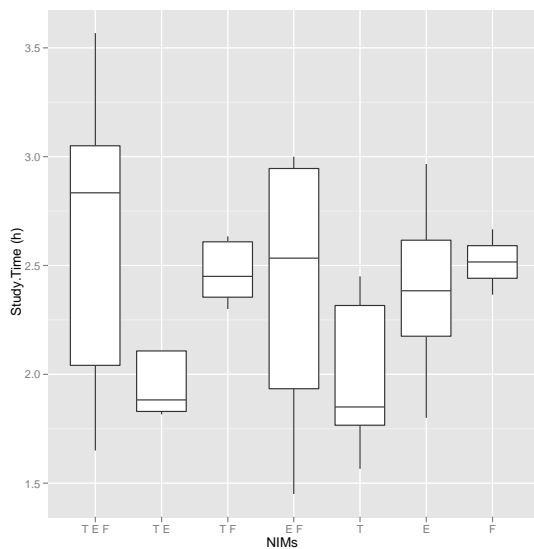


Figure 3(a). Study Time by NIM Set

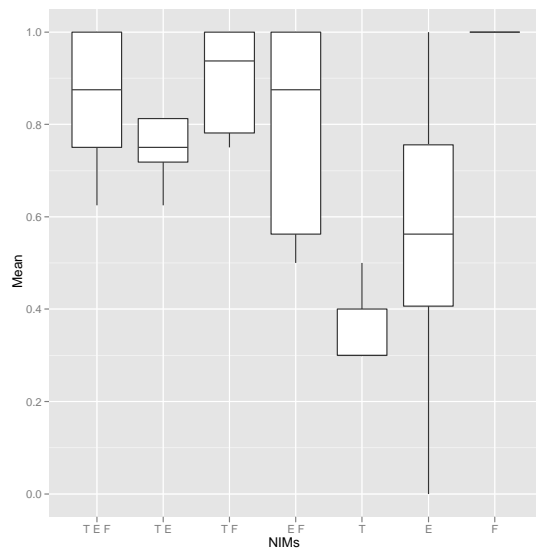


Figure 3(b). Post Test Mean by NIM Set

In Figure 3(b), we show mean post-test score by NIM-set. The double-NIM subjects seemed to do almost as well as those given the full set of NIMs. The subjects given the F NIM seemed to do the best of the restricted sets; in fact, all subjects in the F-only group had a perfect post-test score. There are several reasons why this might be the case. This NIM is somewhat a combination of “by feedback” and “by telling”; in the self-paced “by feedback” lessons, if a subject doesn’t follow the correct procedure, the subject is given a “by telling” description of what should have been done. This is also the only NIM where subjects get any practice with the simulator before the post-test.

We also tried allowing subjects of the lowest performing NIM sets (T and E) to ask questions about the curriculum through a restricted interface. Counter to our expectations, it was very difficult to get subjects to ask anything. Finally, by choosing subjects who knew us and who were majoring in relatively non-technical fields, we got a few questions. Even then, the questions were relatively basic questions about definitions, such as a misunderstanding about the meaning of “absolute value”. Since these subjects did not end up differing very much from the standard groups, we did not separate them in our results.

#### **4.2.2 Qualitative Results**

Creating the context and protocols needed to understand and rigorously evaluate the goals of the Bootstrapped Learning project using human teachers and students has been challenging. We offer the following lessons we have learned in our exploration of this design and empirical space.

Human and electronic students differ in fundamental ways that make it difficult to create analogous contexts without providing one side with undue advantages over the other. First, e-students have perfect memory of all lesson material they have seen. We compensate for this in human testing by allowing subjects to take notes and review lessons if desired. Human students also have a harder time interpreting formal language or concepts expressed in other “unnatural” ways. Because of this, we were forced to produce a more natural version of the e-student curriculum for the human students, introducing possible confounding factors into our comparison.

On the other hand, human students have a greater understanding of the semantics of words and have the ability to gain domain knowledge outside the formal channel of the curriculum, such as through voice intonation or gestures inadvertently expressed by a human teacher. We addressed the issue of “leaky” semantics by being careful that our choice of terms didn’t leak unintentional knowledge and by going through several preliminary iterations of the curriculum. Interestingly, increased semantic understanding was also occasionally detrimental to human subjects, when the knowledge leaked by terms was misleading. The innovation of the self-paced curriculum was critical for addressing the problem of extra-curricular knowledge transfer.

Increased automation of lesson structures was essential for the greater curriculum complexity and greater number of subjects and groups needed for Phase III. For example, we automated the generation of curriculum and test configurations and the subsequent storing of results and grading.

The issue of decreased scores with realtime testing in Phase III was unexpected. We hypothesize it was because of 1) subject boredom and 2) the issue of training vs. education. When testing subjects with realtime simulators, subjects learned that there were states where nothing occurred and would lose focus; we had to be vigilant for the appearance of web-capable smartphones during this time. Additionally, since stricter time-limits were placed on individual tasks (though the time for all tasks together was actually greater), we believe this may be issue of training (skill gained through repetitive practice) rather than education (knowledge gained through learning). Results on models of human task-performance such as the Human Model Processor [Card et al., 1986] and GOMS [John and Kieras, 1996] may be relevant here. We do not know whether this will be an issue for e-students.

## **ACKNOWLEDGEMENT**

This research was sponsored by AFRL under contract FA8650-07-C-7722. Special thanks to our colleagues at BAE Systems, Cycorp, Inc., Sarnoff Corporation, Stottler Henke Associates, Inc., and Teknowledge Corporation for their collaboration in developing the various BL curricula and testing materials, to Matthew Berland for early work on this project, and to study participants at The University of Texas at Austin.

## REFERENCES

- Matthew Berland and Dewayne E. Perry. Novice Human Teachers of a Virtual Toddler: A Case Study. Technical report, The University of Texas at Austin, 2009. URL <http://users.ece.utexas.edu/~perry/work/papers/090123-MB-blexp1.pdf>.
- C. Morrison and D. Bryce and I. Fasel and A. Rebguns. Augmenting Instructable Computing with Planning Technology. In *ICAPS'09 Workshop on the International Competition for Knowledge Engineering in Planning and Scheduling*, 2009.
- S. Card, T. Moran, and A. Newell. The model human processor- An engineering model of human performance. *Handbook of Perception and Human Performance*, 2:45–1, 1986.
- Jon Curtis. BAE Tech Report TR-2231: Bootstrapped Learning Interaction Language. Technical report, BAE, April 2009.
- N. Dahlback, A. Jonsson, and L. Ahrenberg. Wizard of Oz studies – Why and How. *Knowledge-Based Systems*, 6(4):258 – 266, 1993. ISSN 0950-7051. doi: DOI:10.1016/0950-7051(93)90017-N. Special Issue: Intelligent User Interfaces.
- J. Davis, K. Leelawong, K. Belyne, B. Bodenheimer, G. Biswas, N. Vye, and J. Bransford. Intelligent user interface design for teachable agent systems. In *Proceedings of the 8th international conference on Intelligent user interfaces*, pages 26–33. ACM, 2003. ISBN 1581135866.
- A. Frank and A. Asuncion. UCI Machine Learning Repository, University of California, Irvine, School of Information and Computer Sciences, 2010. URL <http://archive.ics.uci.edu/ml>.
- Robert D. Grant, David DeAngelis, Dan Luu, Dewayne E. Perry, and Kathy Ryall. Designing Human Benchmark Experiments for Testing Software Agents. In *Proceedings of the 14th International Conference on Evaluation and Assessment in Software Engineering* (to appear). BCS eWIC, 2011.
- Ralph Grishman and Beth Sundheim. Message Understanding Conference-6: A Brief History. In *Proceedings of the 16th conference on Computational linguistics*, pages 466–471, Morristown, NJ, USA, 1996. Association for Computational Linguistics. doi: <http://dx.doi.org.ezproxy.lib.utexas.edu/10.3115/992628.992709>.
- Bonnie E. John and David E. Kieras. The GOMS Family of User Interface Analysis Techniques: Comparison and Contrast. *ACM Trans. Comput.-Hum. Interact.*, 3(4):320–351, 1996. ISSN 1073-0516. doi: <http://doi.acm.org/10.1145/235833.236054>.
- Dan Oblinger. Bootstrapped Learning: Creating the Electronic Student that Learns from Natural Instruction. AAAI Briefing, 2006. URL [http://www.darpa.mil/ipto/programs/bl/docs/AAAI\\_Briefing.pdf](http://www.darpa.mil/ipto/programs/bl/docs/AAAI_Briefing.pdf).
- R. Rosenthal and R. Rosnow. *Essentials of Behavioral Research: Methods and Data Analysis*. McGraw-Hill Series in Psychology. McGrawHill, New York, second edition, 1991.
- R.K. Yin. *Case Study Research: Design and Methods*. Sage Publications, Inc, 2008.

# QUALITATIVE RESULTS OF AN EXPERIMENT WITH ADJUSTABLE INTERACTIVE RINGS FOR IDTV

Leonardo Cunha de Miranda, Heiko Horst Hornung and M. Cecília C. Baranauskas  
*Institute of Computing, University of Campinas (UNICAMP)*  
*13083-852, Campinas, SP, Brazil*

## ABSTRACT

Literature has indicated that the remote control in its current form is not an appropriate device for mediating the interaction between users and applications for Interactive Digital Television (iDTV). With the completion of a related research project, we proposed a new form of interaction with the iDTV through an interaction language based on gestures that are performed via a new physical artifact of interaction we name Adjustable Interactive Rings (AIRs). This paper presents the qualitative analysis of an experiment conducted with twelve end-users that interacted with a simulated application for iDTV via a hardware prototype of AIRs. This evaluation, among other things, sought to determine the emotional and affective state of participants while using AIRs for iDTV. The results of the qualitative analysis corroborate a positive user acceptance of the AIRs.

## KEYWORDS

Interactive digital television, user experiment, qualitative analysis, hedonic evaluation, gesture-based interaction, human-computer interaction.

## 1. INTRODUCTION

To carry out this work we counted on earlier results – e.g., Miranda *et al.* (2009a, 2009b, 2010) – obtained within the context of designing the interaction of users with applications of Interactive Digital Television (iDTV). From the perspective of Human-Computer Interaction (HCI), our research is directed to physical artifacts of interaction mediating users' interaction with iDTV applications.

There is a consensus in the literature that the remote control in its current form is not the most suitable device to mediate the interaction between users and iDTV applications. Several authors have discussed the problems of user interaction with the remote control, e.g., Nilsen (2004), Berglund (2004), Cesar *et al.* (2008a, 2008b), and Miranda *et al.* (2009b).

The Adjustable Interactive Rings (AIRs) for iDTV are a technology resulting from a research project (Miranda, 2010) which is currently in the final stage of the patenting process. They are physical artifacts of interaction with iDTV with features that distinguish them from other solutions described in literature, for example those presented by Fujita *et al.* (2003) and Sohn and Lee (2004). As specified and implemented by Miranda *et al.* (2010), the resulting physical artifact of interaction is accessible and ambidextrous, and enables flexible use with a simple gesture-based interaction language. During mid-2010 we conducted an experiment with prospective end-users to evaluate the interface and user interaction with the hardware prototype of AIRs for iDTV. Quantitative results of this experiment, i.e. an assessment of users' performance in executing the tasks were presented in Miranda *et al.* (2011).

The aim of this work is to present and discuss qualitative results of this experiment seeking thereby a more refined understanding of how users make sense of, perceive and use the AIRs, i.e. we aim to better understand the relationship of users with the technology we developed. Thus, this paper presents an analysis of this experiment from another point of view than that of Miranda *et al.* (2011), i.e., our focus is directed on the qualitative results of the activity and on exploring the emotional and affective responses of participants to use the AIRs for iDTV.

This paper is organized as follows: in Section 2 we present the AIRs for iDTV as well as the experiment we conducted; in Section 3 we present the theme affective quality in computer systems; in Section 4 we

present a qualitative analysis of the results of the experiment and discuss these results under the perspective of HCI; in Section 5 we conclude and indicate future work.

## 2. THE AIRS FOR IDTV AND THE EXPERIMENT

The AIRs for iDTV are a technology resulting from a research project inspired by the principles of Universal Design (Story, 1998) and conducted taking a participatory approach (Schuler and Namioka, 1993). In our research, we investigated the interaction design in iDTV relying on the cooperation of representatives of the target audience during the entire process of designing and developing the solution; i.e., designing solutions with and for users. In this sense, it is noteworthy that the product design and the interaction language of the artifact were specified together with end-users.

This new digital artifact for iDTV consists of two parts: i) the physical artifact of interaction; and ii) a module interfacing with the receiver (RIM). The physical artifact of interaction comprises three different types of Interactive Adjustable Rings: one for Activation (AIR-A), other for Movement (AIR-M), and a last one for selecting Options (AIR-O). The AIR-M is the only AIR that has its spatial movement captured by the RIM. Each AIR has a single button.

Before implementing the hardware prototype of that new physical artifact of interaction for iDTV with an accessible design and a simple gesture-based interaction language, we performed a series of previous studies, e.g., Miranda *et al.* (2009a, 2009b). After implementing the hardware prototype of the AIRs, we conducted an experiment to analyze the user's interaction with this new device.

The experiment was conducted in July 2010 and was recorded on audio/video with the prior consent of participants. The activity was attended by twelve end-users who had never used the AIRs before. In this paper, these users are identified by U1, U2, ..., U11, U12. The detailed profile of these users is available in Miranda *et al.* (2011). The twelve participants were divided into six pairs, in this paper identified by P1, P2, ..., P5, P6. The constitution of the pairs was as follows: P1(U1, U2) P2(U3, U4), P3(U5, U6), P4(U7, U8), P5(U9, U10), and P6(U11, U12).

The six pairs were divided into two groups: i) pairs without knowledge of AIRs for iDTV; and ii) pairs with basic knowledge on AIRs for iDTV. The pairs belonging to the Group (ii) were P1, P4, and P6. The members of the group without knowledge – group (i) – received no instructions regarding how to use the AIRs.

The twelve users interacted with a simulated application of the iDTV using the AIRs. Our initial plan was that only one participant from each pair would actually use the AIRs. However, all twelve users expressed the desire to use the AIRs, and were allowed to do so. Using the AIRs, the users had to perform six pre-defined tasks: T1) turn on the television; T2) enter name or nickname; T3) move geometric figures; T4) switch from channel 2 to channel 12; T5) adjust the volume; and T6) turn off the television.

The simulated iDTV application had six channels – channels 2, 4, 6, 8, 10, and 12 – circularly shown. The six channels are the channels available on public broadcast television in the region where the experiment was conducted (Campinas, State of São Paulo, Brazil).

At the end of the experiment, the users were asked to anonymously fill in a form with open questions, and to anonymously evaluate the use of AIRs via an instrument that sought to determine the emotional and affective state of participants. Figure 1 shows some moments of the experiment.



Figure 1. Participants using the AIRs during the experiment

### 3. AFFECTIVE QUALITY IN COMPUTER SYSTEMS

The assessment of the affective quality of computer systems is still a challenge. To capture the emotional and affective state of participants using the AIRs prototype, in this study, we used the Self Assessment Manikin (SAM) instrument (Lang, 1985; Morris, 1995; Lang *et al.*, 2005), since it is easy to apply and could be effective regarding our intended purpose.

The SAM form is able to capture three different kinds of emotional and affective states. For each state it uses a line of five core images, allowing an evaluation on a scale of 1 to 9, since it is possible to mark one of the five images or one of the four spaces between two images. The three different kinds of emotional and affective states are:

□ **1<sup>st</sup> line (top):** represents the satisfaction, i.e. if the participant did or did not like the activity. The scale represented by the images goes from happy (left) to sad (right). The middle element can be seen as indifferent, i.e. the participant is neither happy nor sad with the application;

□ **2<sup>nd</sup> line (middle):** represents the motivation, the scale of which is: motivated, excited, awake, bored, sleeping;

□ **3<sup>rd</sup> line (bottom):** represents the control; the scale goes from intimidated/being coerced to being in total control of the situation. Note that this scale is symmetrically inverted with regard to the first two lines.

At the end of each task, the respective participant was asked to fill in the SAM form. As can be seen in Figure 2, the SAM form is strictly iconographic.

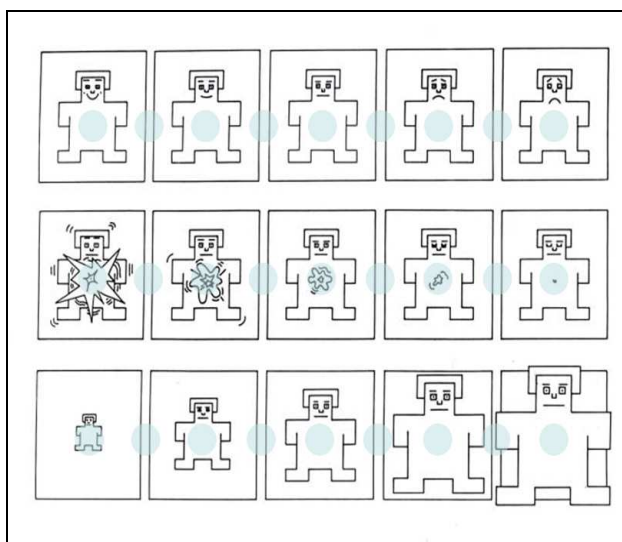


Figure 2. The SAM form

### 4. RESULTS OF THE EXPERIMENT

In this section we first present some qualitative data collected during the experiment in order to then introduce a qualitative analysis and hedonic evaluation of the AIRs.

#### 4.1 Qualitative Observations

Based on the audio/video recordings and data collected by the observer, we point out observations, which later on will be used as input for analysis and discussion of the experiment. The data is grouped by the first five tasks that the participants performed during the experiment. T6 (turn off the television) is omitted, since we did not identify any relevant data to be presented within the context of this work.

##### T1: Turn on the television



□ Initially, U4 tried out different gestures with the AIRs without considering to press the button on any of the AIRs;

□ U6 initially thought that the gesture for turning on the television was pressing the AIR-M and, immediately followed by pressing the AIR-O twice. Later on, during the training session, that user learned that the gesture for turning on/off the TV is to hold the AIR-M for at least two seconds.

**T2: Enter name or nickname**

□ U1 unintentionally turned the TV off and on again;

□ U6 uttered: “If the AIR-M turns the TV on and off, then it shouldn’t do anything else”.

**T3: Move geometric figures**

□ U7 pressed the AIR-M twice.

**T4: Switch from channel 2 to channel 12**

□ U1 went directly from channel 2 to channel 12 “backwards”, i.e., without going through the channels 4, 6, 8, and 10;

□ Although U6 managed to complete the task, considering this participant’s feedback after the experiment, we believe that U6 completed the task without actually being aware of the gestures required for the task;

□ U7 did not complete the task.

**T5: Adjust the volume**

□ From this task on, U5 began to support the arm wearing the AIRs with the other arm;

□ U7 did not complete this task.

**In general: Concerning the use of AIRs by users**

□ U1 has used the left hand to support the right arm. During the experiment this user commented that his/her hand was in pain (not mentioning if it was the right or left hand);

□ U9 and U11 supported the arms on which they wore the AIRs with the other hand (see a similar situation to another user in Figure 1d);

□ U4 said during the experiment: “My arm got tired”;

□ During the execution of all tasks, U12 always pressed the button of the AIRs with the other hand;

□ During the interaction of U6, the partner U5 perceived that the RIM was emitting some kind of feedback on U6’s actions with the AIRs;

□ U9 pointed to the AIR-O and asked the mediator: “Why is this AIR used less than the others?”;

□ U7 thought that in order to change channels or to adjust the volume on the TV he/she had to make gestures relative to the horizontal axis, and that the difference between the two actions was to press or not the button on the AIR-O.

In Table 1 we present on which finger the participants wore each AIR. U11 was the only participant who wore the AIRs with the buttons at the back of the hand. All other participants used the AIRs as we had planned in the conceptual design of the artifact, i.e., with the buttons at the palm.

Table 1. Detailing the use of AIRs by the users during the experiment

User	Pair	AIR-A (green)		AIR-M (yellow)		AIR-O (blue)	
U1	P1	right	index finger	right	middle finger	right	ring finger
U2	P1	right	index finger	right	middle finger	right	ring finger
U3	P2	right	index finger	right	ring finger	right	middle finger
U4	P2	left	middle finger	left	ring finger	left	index finger
U5	P3	right	index finger	right	ring finger	right	middle finger
U6	P3	left	middle finger	left	ring finger	left	index finger
U7	P4	right	index finger	right	middle finger	right	ring finger
U8	P4	right	index finger	right	middle finger	right	ring finger
U9	P5	right	middle finger	right	index finger	right	ring finger
U10	P5	right	middle finger	right	index finger	right	ring finger
U11	P6	left	middle finger	left	ring finger	left	index finger
U12	P6	right	index finger	right	middle finger	right	ring finger

## 4.2 Qualitative Analysis and Hedonic Evaluation

Based on the data collected by the observer (see Table 1), we analyzed how the participants used the AIRs to perform the six tasks. During the experiment, the majority of users (9 out of 12) used the AIRs with the right hand. Figure 3 is a graphical representation of Table 1 grouped by AIR and finger and ignoring the hand.

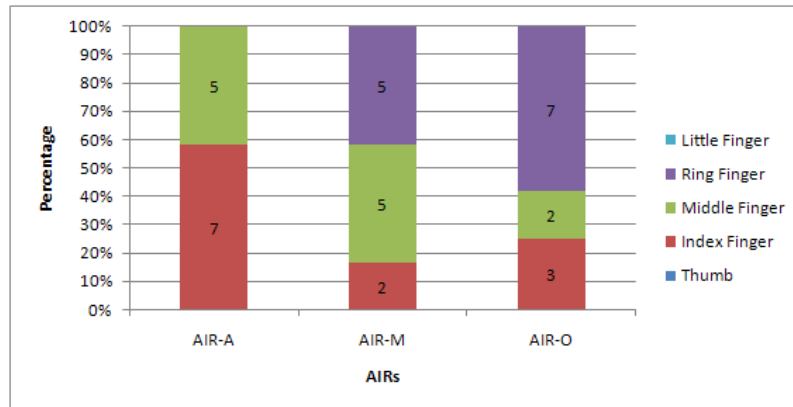


Figure 3. Use of AIRs by fingers

The data presented in Figure 3 gives a clearer picture on how users perceive and use the AIRs. We can infer that there is a tendency to wear the AIR-A (activation) on the index finger and the AIR-O (options) on the ring finger. Moreover, we noted that no user wore any of the AIRs on the thumb or little finger. Perhaps this is due to the anatomy of the human hand and the hardware design of AIRs. The fact that no user wore any AIR on the thumb is important since the AIRs were designed to be used on any finger except for the thumb. It has been envisioned that the thumb would be the digit that could potentially be used to press the single button on each of the AIRs. For detailed information about the product design of the AIRs see Miranda *et al.* (2010).

To analyze the result of the SAM, we adopted the method of assessing the affective quality proposed by Hayashi *et al.* (2008). The method proposed by these authors has its roots in the work of Chorianopoulos and Spinellis (2006), who proposed a framework for assessing user interfaces of iTV applications. The method proposed by Chorianopoulos and Spinellis by its turn was based on Norman (2004). It is noteworthy that this method has been used in other studies which also had as an objective to assess the affective quality in various kinds of digital artifacts, for example, the XO laptop (Silva *et al.*, 2008), games with an interaction design objective (Romani and Baranauskas, 2009), and social software (Piccolo *et al.*, 2010).

Figure 4 shows the accumulated answers of the SAM forms filled in by the 12 participants.

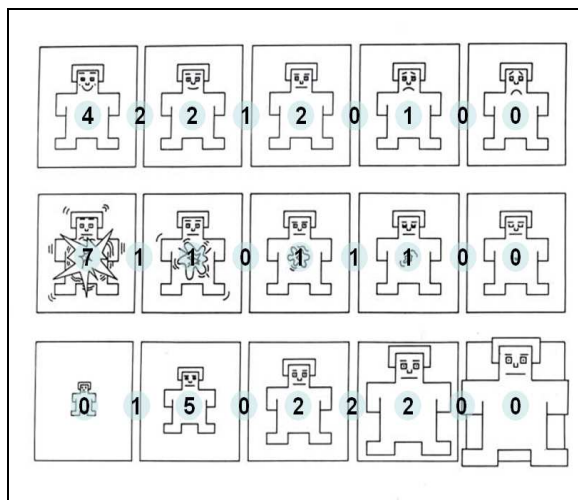


Figure 4. Results of marking the SAM

Table 2 shows the summarized results according to the method proposed by Hayashi *et al.* (2008). Note that the data of the three rows of the SAM form is grouped into three blocks: Negative Ratings (R-), Neutral Ratings (RN), and Positive Ratings (R+).

Table 2. Summarized SAM results

Satisfaction			Motivation			Control		
R-	RN	R+	R-	RN	R+	R-	RN	R+
1	2	9	2	1	9	6	2	4

Thus, based on the summarized result (Table 2), Figure 5 presents the final distribution.

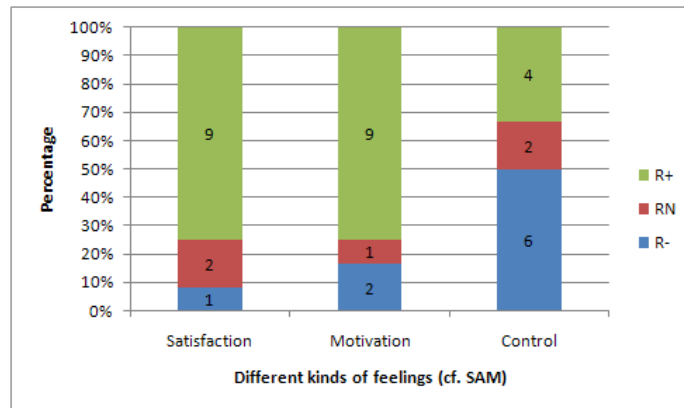


Figure 5. SAM results

From the graph shown in Figure 5 we can see that most participants evaluated the solution positively (see satisfaction and motivation). The only dimension that received negative evaluations was control.

Considering the results presented above, we conclude that the solution had a good reception, because, for example, all participants expressed the desire to use the AIRs, although originally only one participant in each pair experimented the AIRs. Another indicator of satisfaction was discovered during the analysis of the video: all twelve participants smiled while using the AIRs prototype. Some of these moments can be seen in Figure 1. This finding is in concordance with the result obtained by analyzing the SAM forms, where the satisfaction and motivation were markedly positive (Figure 5).

A limitation pointed out by participants was the lack of accuracy of the hardware prototype due to a sensitivity problem of the infrared sensor that did not work as expected during the experiment. This fact was explicitly visible in the result of the SAM analysis – control column in Figure 5 – and some answers given by participants in the evaluation form. Citing one participant: “The feeling of lack of control will be resolved with future versions of the hardware. It was difficult to control the cursor with the AIRs”. It is worth noting that this limitation of the prototype hardware implementation of AIRs has not prevented the experiment of continuation, but in fact it leaves room for improvement.

Except for the problem of the infrared sensor sensitivity, it is worth mentioning the positive responses for the solution, described by users in the evaluation form, for example, the comment from one of the subjects who deemed positive the fact that the AIRs do not have textual labels.

During the experiment we also observed evidence that users wanted to interact with iTV applications using additional forms of interaction, e.g., by using voice command. This became clear when one of the users said: “Turn on!”. This is an important finding towards building a multimodal interface based on Adjustable Interactive Rings.

Finally, the results of qualitative analysis presented in this paper together with quantitative results of the same experiment presented in Miranda *et al.* (2011), confirm the positive user acceptance of the AIRs. During both the quantitative analysis (measuring user’s task performance) and the qualitative analysis (the SAM), the same hardware limitation and the need to improve the sensitivity of the motion-capturing sensor have been identified. It is also worth mentioning that the qualitative analysis presented in this paper, besides presenting additional results regarding the experiment, also helps to better understand the outcome of the quantitative analysis presented previously in Miranda *et al.* (2011).

## 5. CONCLUSION

This paper presented results of an experiment with the hardware prototype of the Adjustable Interactive Rings for iDTV. A qualitative analysis of the interface and interaction between users and AIRs presented in this work corroborate previous results on the positive user acceptance of the AIRs. The obtained results complement quantitative results presented previously in Miranda *et al.* (2011).

These results help us to advance in our research on physical artifacts of interaction for iDTV to consider a multimodal interface for iDTV based on AIRs for other contexts. Further work involves to elicit requirements for the multimodal interface based on AIRs that we aim to design and implement.

Despite the fact that AIRs have been developed for the use with iDTV, we believe that the positive results suggest that this technology could be used as a technological basis in environments other than iDTV. For example, one area we are envisioning is the use of AIRs as interaction artifact in the context of computer games.

## ACKNOWLEDGEMENTS

This work was supported in part by the State of São Paulo Research Foundation – FAPESP (grant #2010/11004-9), by the National Council for Scientific and Technological Development – CNPq (grants #141489/2008-1 and #141058/2010-2), and by the Brazilian Federal Agency for Support and Evaluation of Graduate Education – CAPES (grants #01-P-09881/2007 and #01-P-08503/2008).

## REFERENCES

- Berglund, A., 2004. *Augmenting the Remote Control: Studies in Complex Information Navigation for Digital TV*. Doctoral Thesis. Linköping University.
- Cesar, P., Bulterman, D.C. and Jansen, A.J., 2008a. Usages of the Secondary Screen in an Interactive Television Environment: Control, Enrich, Share, and Transfer Television Content. *Proceedings of the 6<sup>th</sup> European Conference on Interactive Television*. Salzburg, Austria, pp. 168–177.
- Cesar, P., Chorianopoulos, K. and Jensen, J.F., 2008b. Social Television and User Interaction. *ACM Computers in Entertainment*, Vol. 6, No. 1, pp. 1–10.
- Chorianopoulos, K. and Spinellis, D., 2006. User Interface Evaluation of Interactive TV: A Media Studies Perspective. *Universal Access in the Information Society*, Vol. 5, No. 2, pp. 209–218.
- Fujita, K., Kuwano, H., Tsuzuki, T., Ono, Y. and Ishihara, T., 2003. A New Digital TV Interface Employing Speech Recognition. *IEEE Transactions on Consumer Electronics*, Vol. 49, No. 3, pp. 765–769.
- Hayashi, E.C.S., Neris, V.P.A., Martins, M.C., Baranauskas, M.C.C., Piccolo, L.S.G. and Costa, R., 2008. Avaliando a Qualidade Afetiva de Sistemas Computacionais Interativos no Cenário Brasileiro. In: A.M. Melo, L.S.G. Piccolo, I.M.A. Avila and C.A. Tambascia, eds. 2009. *Usabilidade, Acessibilidade e Inteligibilidade Aplicadas em Interfaces para Analfabetos, Idosos e Pessoas com Deficiência*. Fundação CPqD, Campinas, Brazil, pp. 55–62.
- Lang, P.J., 1985. The Cognitive Psychophysiology of Emotion: Fear and Anxiety. In: A.H. Tuma and J.D. Maser, eds. *Anxiety and the Anxiety Disorders*. Lawrence Erlbaum Associates, Inc., Hillsdale, United States, pp. 131–170.
- Lang, P.J., Bradley, M.M. and Cuthbert, B.N., 2005. *International Affective Picture System (IAPS): Instruction Manual and Affective Ratings* (Technical Report A-6) The Center for Research in Psychophysiology, University of Florida.
- Miranda, L.C., 2010. *Artefatos e Linguagens de Interação com Sistemas Digitais Contemporâneos: Os Anéis Interativos Ajustáveis para a Televisão Digital Interativa*. Doctoral Thesis. University of Campinas.
- Miranda, L.C., Hayashi, E.C.S. and Baranauskas, M.C.C., 2009a. Identifying Interaction Barriers in the Use of Remote Controls. *Proceedings of the 7<sup>th</sup> Latin American Web Congress*. Merida, Mexico, pp. 97–104.
- Miranda, L.C., Hornung, H.H. and Baranauskas, M.C.C., 2009b. Prospecting a Gesture Based Interaction Model for iDTV. *Proceedings of the IADIS International Conference on Interfaces and Human Computer Interaction*. Algarve, Portugal, pp. 19–26.
- Miranda, L.C., Hornung, H.H. and Baranauskas, M.C.C., 2010. Adjustable Interactive Rings for iDTV. *IEEE Transactions on Consumer Electronics*, Vol. 56, No. 3, pp. 1988–1996.

- Miranda, L.C., Hornung, H.H. and Baranauskas, M.C.C., 2011. Adjustable Interactive Rings for iDTV: First Results of an Experiment with End-Users. *Proceedings of the 2<sup>nd</sup> International Conference on Human Centered Design*. Orlando, United States, pp. 1–10, in press.
- Morris, J.D., 1995. Observations: SAM: The Self Assessment Manikin: An Efficient Cross-Cultural Measurement of Emotional Response. *Journal of Advertisement Research*, Vol. 35, No. 6, pp. 63–68.
- Nielsen, J., 2004. *Remote Control Anarchy*. [Online] Available at: <http://www.useit.com/alertbox/20040607.html> [Accessed 29 March 2007].
- Norman, D.A., 2004. *Emotional Design: Why We Love (or Hate) Everyday Things*. Basic Books, New York, United States.
- Piccolo, L.S.G., Hayashi, E.C.S. and Baranauskas, M.C.C., 2010. The Evaluation of Affective Quality in Social Software: Preliminary Thoughts. *Extended Proceedings of the IX Brazilian Symposium on Human Factors in Computing Systems*. Belo Horizonte, Brazil, pp. 29–38.
- Romani, R. and Baranauskas, M.C.C., 2009. GWIDO - Games with Interaction Design Objective. *Proceedings of the IADIS International Conference on WWW/Internet*. Rome, Italy, pp. 351–358.
- Schuler, D. and Namioka, A., 1993. *Participatory Design: Principles and Practices*. Lawrence Erlbaum Associates, Inc., Hillsdale, United States.
- Silva, F.B., Romani, R. and Baranauskas, M.C.C., 2008. SOO Brasileiro: Aprendizagem e Diversão no XO. *Revista Brasileira de Informática na Educação*, Vol. 16, No. 3, pp. 29–41.
- Sohn, M. and Lee, G., 2004. SonarPen: An Ultrasonic Pointing Device for an Interactive TV. *IEEE Transactions on Consumer Electronics*, Vol. 50, No. 2, pp. 413–419.
- Story, M.F., 1998. Maximizing Usability: The Principles of Universal Design. *Assistive Technology*, Vol. 6, No. 1, pp. 4–12.

# DESIGN CREATION SYSTEM BASED ON DESIGNER'S KANSEI AND CONSUMER'S PREFERENCE

Hideki Aoyama and Kentaro Ohshima

*Keio University, Department of System Design Engineering  
3-14-1 Hiyoshi, Kohokuku, Yokohama, 223-8522 Japan*

## ABSTRACT

Making product differentiation in quality has been becoming difficult with progress of production technology, and aesthetic design is given increasing importance in the charm of products. Previously, appearance design was creation activity strongly depending on the aesthetic sense of a designer. Nowadays, it is demanded to survey consumer's needs on appearance design and reflect them on product design; in other words, design based on only designer's Kansei is not enough for customer's preference. It becomes important challenge to effectively gather consumer's needs, extract valuable information from the needs and apply them for product development. There are several works which create a shape in harmony with consumer's preference. However, they do not accept designer's intention in the design process. This study thus aims to construct a design support system which can provide design ideas (design solutions) to the designer at an early stage in the design process. The provided design ideas are generated from a designer's original idea by adding consumer's preference. A system to collect consumer's preference data using interactive genetic algorithm was constructed. The system can make if-then rules as consumer's preference based on rough sets theory and create design solutions by applying the consumer's preference rules on a designer's idea.

## KEYWORDS

Design creation system, Kansei, consumer's need, consumer's preference, interactive genetic algorithm, rough set theory

## 1. INTRODUCTION

Since making product differentiation in quality has been becoming difficult with progress of production technology, contribution of functions and appearances in product design are increasing in product charm. Up to now, appearance design has been conducted by designer's creativity which depends on aesthetic Kansei of a designer. However, nowadays, it is strongly desired to make a product shape by not only designer's Kansei but also consumer's preference. Therefore, it is becoming important to survey and analyze consumer's preference and to reflect them on product design. In the reflection of consumer's preference, designer's idea/Kansei should be fundamental. There are several works<sup>1), 2), 3)</sup> which create a shape according to consumer's preferences based on scientific methods. Generally, QFD (Quality Function Development) is an effective tool for extracting consumer's preferences. QFD systematically develops technical elements by analyzing the data of questionnaire on product requirements. However, the analysis results depend on the setting of requirement issues and scoring of technical elements. Since QFD processes requirements represented with language, it is not suitable for analyzing technical elements based on Kansei such as design and operability<sup>2)</sup>. Besides, the previous works cannot make simultaneous use of both consumer's preferences and designer's Kansei.

Thus, in this research, in order to support early stage of product design, a system which survey consumer's preferences with easy operations, automatically extract preference rules by analyzing survey results, and automatically create several design solutions by adding the consumer's preferences rules on a designer's design solution is developed. In other words, the system can automatically create many design solutions by evolving a designer's idea with consumer's preferences. Creating product designs based on both designer's idea and consumer's preferences is the unique and contribution point of this research.

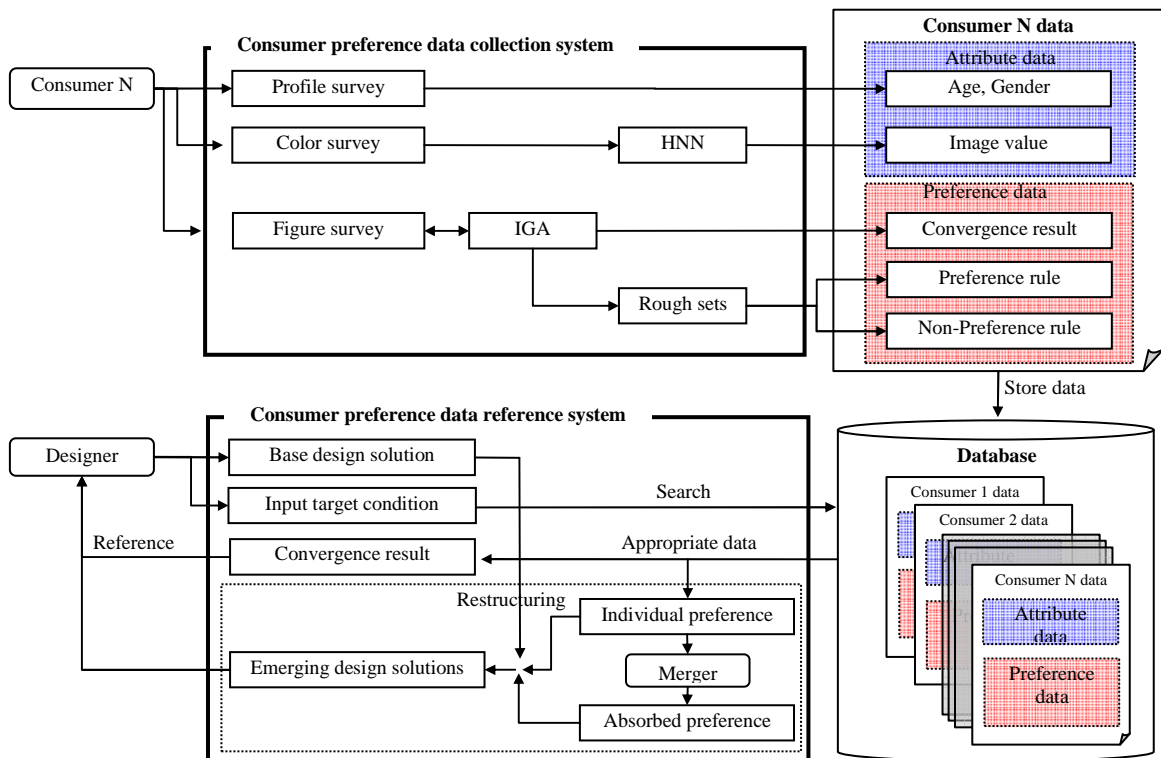


Figure 1. Outline of the developed system

## 2. DESIGN CREATION METHOD BASED ON CONSUMER'S PREFERENCES

### 2.1 Outline of Design Creation Method

The developed system is composed of the two systems; one is a consumer preference data collection system which collects and analyzes customer's preferences on product appearances and the other is a consumer preference data reference system which creates design solutions based on both of designer's idea and customer's preferences. Figure 1 shows the utilization procedures of the system.

In using of the consumer preference data collection system, customers input preference information by selecting favorite designs from design solutions which were automatically created by interactive genetic algorithm (IGA). The collected information is analyzed by rough sets theory and customer's preference rules are then extracted. In the next step, cubes with various colors are indicated and customers select a cube having the favorite color. The selected result (color) is analyzed by a system constructed by hierarchical neural network (HNN) and the impression of the selected color which means customer's requirement image for a product is derived. The survey results on shape and color are stored in the database with the customer profile.

In using of the consumer preference data reference system, the preference rules of targeted customers are extracted from the database constructed by the consumer preference data collection system. Several design solutions called emerged designs are automatically generated by evolving a designer's idea solution with the extracted customer's preferences rules. The emerged designs (design solutions) are indicated to the designer for re-planning the design from the view point of customer's preferences.

## 2.2 Method to collect Customer's Preferences Data and Extract Customer's Preferences Rules

Customers do not have concrete idea on preferences for a target product before purchase of it so that they select favorite products in making comparison-evaluation of each product in target products. In other words, customers make their preferences obvious and concrete through the process of selecting favorite products from the target products<sup>2)</sup>. Therefore, it must be an effective method to collect evaluation data: like or dislike, for design shapes and to extract preference rules by analyzing the data. However, it is not a realistic method to get evaluation data for huge number of design solutions generated by changing all design parameters defining a target design shape. For this problem, in this study, interactive genetic algorithm is applied to reduce the load of consumers in collecting data of customer's preferences. However, in applying interactive genetic algorithm to evaluation based on human sense, there are problems to be solved as follows.

(1) Reduction of the number of design solutions to be evaluated and correspondence for plural preferences

When design parameters of a target product shape are genes, the number of individuals generated by combinations of the all genes must be huge<sup>4)</sup>. It is unrealistic to collect evaluation data for all individuals. It is important to make the rational number of individuals for collecting customer's preferences. In this study, a function to predict evaluation of design solutions based on few evaluation results of customers is introduced. The prediction function is improved as the collected customer evaluation data is increased. In the first generation, 560 target product shapes are created by random combinations of genes, and 24 dissimilar individuals are indicated to customers for evaluation. In the next generation, 12 product shapes which are highly evaluated by the prediction function and 12 product shapes which are dissimilar product shapes for highly evaluated product shapes are indicated for evaluation. These processes are repeated until the solutions of interactive evolutionary computation (IEC) are converged. The evaluation process using 12 highly evaluated product shapes and 12 dissimilar product shapes for highly evaluated product shapes can correspond to plural preferences of human.

(2) Definition of prediction function

In order to develop the prediction function, it is necessary to identify shape features which customers pay attention to. However, identifying the shape features is difficult in little evaluation data (preference data). Therefore, in this study, the prediction function was developed using dissimilarity which represents the difference degree between two design solutions. In the recognition process, shape recognition is smaller effect than impression recognition for the individual difference. It is highly desired that a function to derive the common dissimilarity with highly matching among many customers is obtained. The dissimilarity is calculated by the Euclidean distance with the weight which is derived by the analytic hierarchy process (ARP).

(3) Correspondence to plural preferences

Interactive genetic algorithm (IGA) can find the highest preference. However, there is a possibility that a human has plural preferences. In this case, the second preference and the third preference are neglected. Therefore, in this study, parallel genetic algorithm (PGA) based on analysis of evaluation tendencies of design solutions is applied for corresponding plural preferences.

(4) Preference rules

Interactive evolutionary computation (IEC) can find the best solution in a local solution area, but it is difficult to grasp the preference tendencies in all area of the design solutions. Therefore, in this study, in order to grasp the tendencies in all area of the design solutions, preference data is analyzed by rough sets theory and the causal correlation between shape characteristics and evaluations of design solutions is extracted as if-then rules.

(5) Identification of demand impression (image) for a product

To identify the customer demand of impression (image) for a product is indispensable to search customer's preferences. Since identification of demand images using impression words on product shapes is difficult, this study analogized customer demands for a product by the causal correlation between colors and impressions.



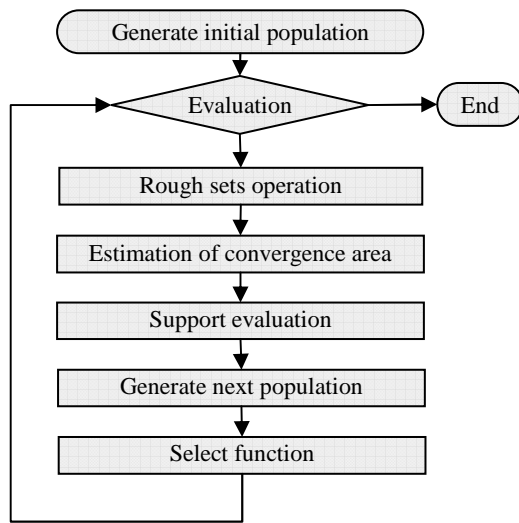


Figure 2. Procedure of collection of preference data on form

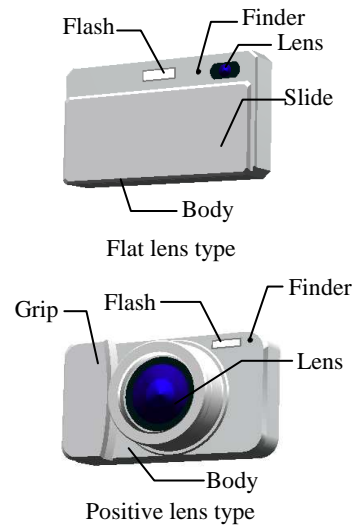


Figure 3. Camera design

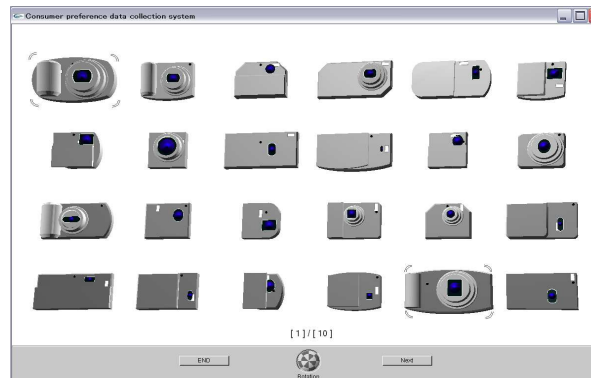


Figure 4. Interface for figure survey

## 2.3 Consumer Preference Data Collection System

A consumer preference data collection system was developed based on the above mentioned method. In the following, collection of preference data on shape, extraction of preference rules, and analogy of impression from color are described.

### 2.3.1 Collection of Preference Data on Form

Effective preference manifestation can be executed by collecting customer's preference data on individuals (design shapes) which are generated by reflecting customer's preferences on the shapes according to the procedures shown in Fig.2. Each operation is as follows.

(1) Generation of first individuals (first design shapes)

By incorporating the shape parameters of digital camera bodies shown in Fig. 3 into gene elements of genetic algorithm, the first individuals (first design shapes) are generated by the random combinations of the shape parameters. The searched area in genetic algorithm is widely kept by selecting individuals (design shapes) in order of dissimilarity. This system generates 560 individuals in a generation, and indicates 240 individuals to customers from them. In order to collect customer preferences, as shown in Fig. 4, 24 individuals among the 240 individuals are simultaneously displayed on a screen and a customer points out the individuals which he/she likes.

(2) Rough sets operation<sup>5)</sup>

Common shape characteristics in the individuals pointed out by customers as like individuals are extracted by rough sets operation and the extracted characteristics are constructed as preference rules.

(3) Correspondence for plural convergence areas

The convergence areas of generic algorithm are estimated by executing cluster analysis of individuals (design solutions) pointed out by customers and recognizing the clusters more than a threshold as convergence areas. For the plural convergence areas, a gene is divided into parts in gene manipulation by parallelization of genetic algorithm and the following processes (4) – (6) are repeated.

(4) Estimation of evaluation of individuals which are not evaluated by customers

The evaluation function based on dissimilarity for each convergence area is developed to estimate the evaluation of individuals which are not evaluated by customers. The estimation of the evaluation is carried out by the evaluation function. Five individuals which are highly evaluated by the evaluation function are selected. This system adds the five individuals into parent individuals to get many evaluation results without customer's real evaluation.

(5) Creation of next generation individuals<sup>(6), 7)</sup>

In this system, individuals are created by two ways; one is based on genetic algorithm and the other is based on preference rules. In the creation of individuals by genetic algorithm, parent individuals are extracted from each cluster divided by roulette selection, and an offspring is created from two extracted parents by uniform crossover. In the uniform crossover, conditions which make shape consistency are not applied. The mutation rate is variable to make convergence and diversity. On the other hand, in the creation of individuals by preference rules, individuals are created based on the shapes defined by preference rules extracted by rough sets theory to an individual created by shape parameters given by a random function.

(6) Extraction of individuals to be indicated to customers

The individuals which are indicated by customers for collecting preference data are determined in order of the priority level which is derived by the covering index function, the solution search function, and the area search function.

(i) Covering index function

The covering index function is an evaluation function to extract preference rules in order of effectiveness. The effectiveness is determined with C. I. (covering index) value.

(ii) Solution search function

The solution search function is an evaluation function to support convergence in genetic algorithm. The dissimilarity between created individuals and individuals which are not indicated to customers for collecting preference data is derived as the distance between the individuals. Individuals are indicated in order of long distance to not indicated individuals.

(iii) Area search function

The area search function is an evaluation function to extend searching area of individuals. The dissimilarity between created individuals and individuals which are indicated to customers for collecting preference data are derived as the distance between the individuals. Individuals are indicated according to the dissimilarity.

### 2.3.2 Extraction of Preference Rules

Several shapes which customers prefer are collected by the system using genetic algorithm. In order to understand the whole preference, causal relationships between shape characteristics and preference evaluation are derived by rough sets theory as customer's preference rules. The customer's preference rules enable a designer to design shapes which customers are satisfied with. Design shapes which are preferred by many customers can be obtained using the rules which are combined with preference rules of many target customers.

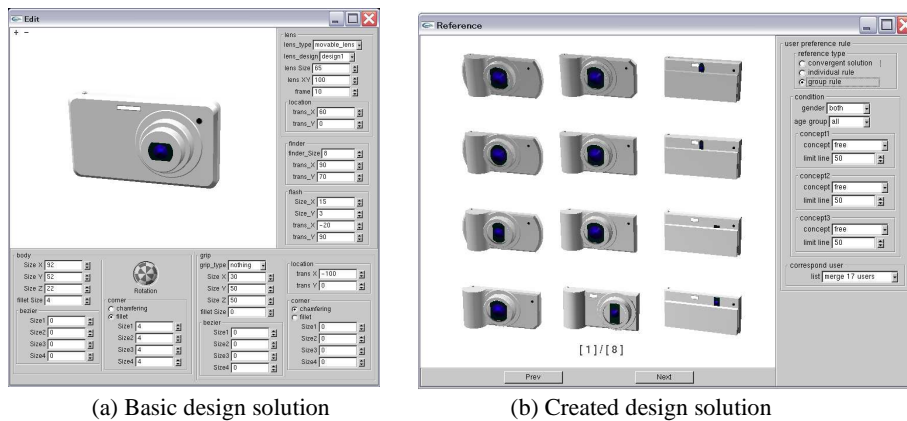


Figure 5. Consumer preference data reference system

### 2.3.3 Estimation of Product Image (Impression)

The product image (impression) which a customer demands of a product is analogized by causal relationships between colors and images<sup>8), 9)</sup>. A customer selects a cube with a color from 256 cubes with different colors. The impression which is represented by 10 image words with weights: showy, cute, clear, individual, feminine, elegant, elegant, earnest, friendly, and relaxing, are estimated from the HSV value of the selected cubic color by a hierarchical neural network (HNN). The learning data for the HNN was collected by questionnaires.

## 2.4 Utilization Method of Preference Data

The preference rules are constructed as if-then format so that they are not suitable for intuitive use by designers. Therefore, in this system, as shown Fig. 5, several potential solutions which are created by applying customer's preference rules to a design solution (basic design solution) designed by a designer are shown. Many design solutions can be created from a basic design solution by adding customer's preference rules. A designer can visually grasp many design solutions which have both Kansei of a designer and customers. This supports the process in which a designer makes concrete ideas from a concept image. By using preference rules of limited customers, design solutions for only targeting customers can be generated.

## 2.5 Reference System of Preference Data

This system creates design solutions according to the following procedures.

- (1) Creation of basic design solution (designer solution)

A designer constructs a designer solution by operating shape parameters using the interface shown in Fig. 6(a).

- (2) Retrieval of preference rules

A designer retrieves customer's preference rules by giving information of target customers: age, sex, and product impression.

- (3) Presentation of created design solutions

As shown in Fig. 6(b), the system presents the design solutions which are created by adding retrieved preference rules to a basic design solution. A designer can reconsider the design by referring the created design solutions which reflect both Kansei of a designer and targeted customers.

### 3. EXPERIMENTS OF SYSTEM EVALUATION

Table 1. Evaluation difference with design base solution

	design base solution	individual	group rule												
			non-segment	gender		concept image									
				male	female	fleshy	cute	clear	unique	feminine	elegant	accomplished	dismissive	domesticable	soft
Average of evaluation value	4.94	8.18	5.41	5.12	6.47	4.94	5.24	3.53	5.12	5.53	5.65	6.00	5.88	5.59	6.18
Average of difference	-	3.24	0.47	0.18	1.53	0.00	0.29	-1.41	0.18	0.59	0.71	1.06	0.94	0.65	1.24
Sample variance	-	6.89	2.96	4.97	2.01	4.35	7.38	8.36	8.26	8.60	7.03	12.17	8.76	8.11	2.89
Unbiased variance	-	7.32	3.14	5.28	2.14	4.63	7.85	8.88	8.78	9.13	7.47	12.93	9.31	8.62	3.07
SEM of difference	-	0.66	0.43	0.56	0.35	0.52	0.68	0.72	0.72	0.73	0.66	0.87	0.74	0.71	0.42
statistics <i>t</i>	-	4.93	1.10	0.32	4.31	0.00	0.43	-1.95	0.25	0.80	1.06	1.21	1.27	0.91	2.91
Significant level 5.0%		OK			OK										OK
Significant level 1.0%		OK			OK										

The effectiveness of the developed system was verified by experiments that customers evaluated the design solutions created by the system.

#### 3.1 Experiment Procedures

In order to verify the efficient collection of preference data and effectiveness of the extracted preference rules, the following procedures are carried out for 17 subjects (12 males and 5 females) as verification experiments.

(1) Customer's preference rules were obtained from convergence solutions in GA by collecting preference data of 17 subjects using the developed preference data collection system.

(2) Design solutions were created by reflecting the customer's preference rules on a basic design solution using the customer preference data reference system. In the experiment, the basic design solution was modeled by refereeing a commercial digital camera. Individual preference rules and group preference rules were used to create design solutions. The group preference data had classified into the groups segmented by sex, age, and concept conditions.

(3) The basic design solution and the created design solutions were evaluated with 10 steps in the experiments.

#### 3.2 Experiment Results

Table 1 shows the results of the evaluation experiments with 10 steps. It is understandable that the design solutions reflecting the preference rules on the basic design solution have higher evaluation points than the basic design solution. In order to verify the effectiveness of the increase of the evaluation points, *t*-tests were carried out under the condition of Gauss distribution of population. The *t* values of the significant levels 5 % and 1 % were 2.12 and 2.29, respectively, for 17 samples and 16 freedom degrees. In the case of larger *t* values than these *t* values: 2.12 and 2.29, it can be understandable that the increase of the evaluation points was beyond the range of accidental errors. Table 1 shows the statistics *t* values calculated from the evaluation differences between the basic design solution and the created design solutions. In Table 1, "OK" means that the design has significance. In the results, it was confirmed that the proposed system could

generate design solutions suiting customer's preferences for the individual, female, and soft concept cases. However, the effectiveness should be verified for more many samples. In the experiments, the indicated population was 6.3 generations in the average, and the time to finish the evaluation was about 3 minutes. From these points, it is also understandable that the proposed system is could extract effective preference rules.

## 4. CONCLUSIONS

This study proposed a design support system which reflects customer's preferences on a design solution created by a designer and create many design solutions suiting customer's preferences. The results are summarized as follows.

(1) A design support system which is composed of a consumer preference data collection system and a consumer preference data reference system was developed.

(2) The consumer preference data collection system can collect customer's preference data by interactive genetic algorithm and can extract the preference rules as if-then rules by the rough sets theory.

(3) The consumer preference data collection system can work well for plural convergences of preference by dividing a gene into parts according to the convergence situations.

(4) The consumer preference data reference system can generate many design solutions suiting customer's preferences by refereeing preference rules on a design solution created by a designer.

(5) A designer can refer wide design solutions by refereeing the design solutions generated by the system, which support making concrete ideas of a designer.

(6) The effectiveness of the developed system was verified by evaluation experiments for a design.

## REFERENCES

- 1) Gershenson J A, Khadilkar D V, and Stauffer L A, Organizing and managing customer requirements during the product definition phase of design, ASME Des Eng Div, Vol.68 (1994), pp.99-107.
- 2) Yanagisawa H, Murakami T, Fukuda S, 2006, Favored-Feature-Based Shape Generation – Externalization of Latent Preference -, proceedings of DETC'06/ASME International 24<sup>th</sup> Computer and Information in Engineering, USA, CD-ROM DETC2006-99136, pp. 1-10.
- 3) H. Yanagisawa, V. Kostov, S. Fukuda, Development of Industrial Design Support System Using Impression Words : With Attention Paid to Diversity of Kansei, Journal of Japan Society of Mechanical Engineers, C Vol. 67-657 (2001), pp. 1682-1688. (in Japanese)
- 4) Zhenyu Gu, Ming Xi Tang, John Hamilton Frazer, 2006, Capturing aesthetic intention during interactive evolution, Computer-aided design, No.38, pp.224-237.
- 5) Pawlak Z, Rough Sets, Int. J. Computer and Information Science, Vol.11, No.5(1982), pp.341-356.
- 6) Li Qiang, Yoichiro M, 2009, Fuzzy Adaptive Search Method for Parallel Genetic Algorithm Based on Evolution History, Biomedical fuzzy and human sciences the official journal of the Biomedical Fuzzy Systems Association , Vol. 1, pp.123-130.
- 7) TAKAGI H, OHYA K, OHSAKI M, 1996, Improvement of Input Interface for Interactive Genetic Algorithms and its Evaluation, Methodol Concept Des Appl Intell Syst, Vol. 2, pp.490-493.
- 8) KIM J, LEE J, CHOI D, 2003, Designing emotionally evocative homepages: an empirical study of the quantitative relations between design factors and emotional dimensions, Int J Hum Comput Stud, Vol.59, No.6, pp. 899-940.
- 9) KIM So YeongLEE, Seung Hee, 2009, CATEGORY METHOD PROPOSAL OF COLOR KANSEI INFORMATION WHICH USES MPI, KANSEI Eng Int, Vol.8, No.2, pp.165-168.

# COUPLED MOBILE DISPLAYS FOR CREATIVITY TECHNIQUE BASED PROBLEM SOLVING PROCESSES IN TABLETOP ENVIRONMENTS

Marc René Frieß and Martin Kleinhans  
*Technische Universitaet Muenchen  
 Boltzmannstraße 3, 85748 Garching*

## ABSTRACT

Single display groupware (SDG) provides an adequate IT environment for co-located creativity support. Despite their potential benefits in regard to communication, coordination and interpretation, such environments also encompass certain disadvantages. Coupled displays are a means to solve some of those disadvantages by enhancing the public SDG workspaces with parallel and private input channels. It's our aim within this paper to present such an approach. After discussing the advantages and disadvantages of creativity support in SDG scenarios, we will review related work on coupled displays in general. In a next step we will present our scenario, a multi-touch tabletop application supporting collaborative creativity techniques. In a next step we will present a concrete implementation based on iPad devices. Finally we will describe an evaluation conducted in a university setting and discuss the corresponding results and implications for further work.

## KEYWORDS

Multi Display Environment, Tabletop, Coupled Displays, Mobile Computing, Creativity Support System

## 1. INTRODUCTION

Single display groupware (SDG) provides a novel way of IT support for co-present collaboration. While in traditional IT environments (e.g. single user PCs) the physical layout of technology can inhibit the interactions between the people involved, SDG allows for novel interaction paradigms which combine the general advantages of IT-support (such as permanent recording and sustained and parallel manipulability of collaborative artifacts) with the advantages of added face-to-face collaboration such as body language and the immediacy of verbal and nonverbal communication (Hilliges2007). This leads to an increased visibility of action and group awareness. When every group member has the possibility to understand at any time what other members are doing, the isolation of single individuals is avoided (Hilliges2007). Novel SDG devices based on multi-touch tabletop displays even intensify the regarded benefits: By processing multiple touch points at the same time they enable a concurrent control of an application by directly interacting with the shared display in an intuitive way.

Especially in the field of creative problem solving, with its core requirements of *communication*, *coordination* and *interpretation*, such novel workspaces are promising (Hilliges2007). However, research on creativity also remarks that during early phases of the idea generation process, group collaboration can be problematic. For example a study on the effectiveness of the (group) creativity technique Brainstorming showed, that nominal groups are more effective (in both quantity and quality of ideas) than groups where participants work collaboratively (Taylor1958). This can be explained by three factors (Isaksen1998):

1. **Group pressure** can result from fear of judgment by others and power imbalances in the group. It inhibits participation and can lead to unwanted conformity of idea proposals.

2. **Social loafing** describes the tendency of group members to do less than their potential would allow them to do. It can occur either if a group member feels isolated from the group or if he feels too submerged.

3. **Production blocking** is regarded as the dominant factor for efficiency losses in group brainstorming processes. Production blocking refers to the fact, that in interacting groups only one member can speak at a

time, while the others have to listen; hence all but one member of the group are blocked and cannot work on their own ideas in this time.

Traditional (distributed) IT support based on personal computers can help to mitigate those effects by e.g. providing anonymity (to avoid group pressure) and parallel input channels (to avoid production blocking). However, those benefits are gained at the expense of the acceptance of the IT system (Magerkurth2001) and the satisfaction of the participants (Dennis2004, p.1), who “*may not be primarily concerned with the number of ideas generated when planning a brainstorming session, but rather may equally desire group well-being and member support*”. Also later phases of a creative process (or different creativity techniques) may require mutual stimulation within the group to e.g. broaden the idea horizon (Magerkurth2001). These (group) aspects are typically supported by SDG environments in a more optimal way, as already discussed before.

To bring together the advantages of SDG and traditional (distributed) IT support for creativity one could regard “real world” scenarios. There, people are free to reference private notes or make private sketches, while collaborating around a public artifact. At any time, people can choose to make private information public or public information private (Shoemaker2000). Multi display environments (MDEs) combining a shared SDG device (such as a tabletop display) and smaller coupled displays (such as smartphones or tablet-PCs) are one possible solution to allow for a similar style of interaction. While on the shared display all information is by default public, the coupled displays allow the participants to also have private information spaces available. As pointed out by Magerkurth et al. (2010), combining individual and group work in this way can lead to more optimal results.

This paper will regard how an application supporting collaborative creativity techniques on a multi-touch tabletop display can be extended by such coupled displays. Therefore, in a first step, related work about coupled displays in general will be regarded. Derived from this related work but also taking into account the theoretical background of an existing tabletop based creativity support system, we will discuss conceptual considerations and describe the resulting implementation. In a next step, an evaluation of this application will be presented. The paper will conclude with a discussion about our findings as well as future work.

## 2. RELATED WORK

Several research activities on coupled displays in single display groupware scenarios have been conducted within the last years. Magerkurth et al. (2004) proposed the combination of a digital tabletop display and personal digital assistants (PDAs) for board games. While the tabletop acted as public space representing the game board and detecting tangible objects on it, the PDAs were used for displaying private information about the game’s artifacts. Especially for board games which involve competition and cooperation, a private communication channel between players was integrated. In a similar scenario, Echter et al. (2009) regarded the use of coupled displays for collaboratively solving Sudoku puzzles at a multi-touch tabletop display, at PDAs and in a mixed (tabletop and PDA) scenario. In their application, they presented the same view on the Sudoku game board on the PDAs as well as on the tabletop and did not differentiate between private and public information.

Magerkurth and Prante (2001) regarded the use of PDAs for creativity support. The PDAs were used as input device for co-located creativity sessions in front of a large virtual wall, but also for individual idea collection outside the (direct) temporal and spatial context of the creativity session. In this regard it was emphasized that ideas can emerge everywhere and every time, especially if someone is not focused directly on solving a problem. Magerkurth’s coupled display application includes two views, one for creating/editing an idea by drawing sketches, inserting texts and assigning a title, while in the other, ideas can be clustered on a virtual pin board. Such a cluster can then be transmitted to a larger virtual wall to present the ideas to a broader audience.

The work by Greenberg et al. (1998, 1999) also targeted the combination of large (public) and small (private) displays. However the main focus of their research lies on the specific characteristics of public and private information and the transitions between them. In their work they discuss problems resulting from an interface-dependent design of interaction objects (e.g. having textual representations on the PDA and images in the large SDG display). In this regard, Magerkurth and Prante (2001) propose to use identical representations of interaction objects throughout different devices/interfaces. In another study, Carter et al. (2004, p.1207) regarded “*methods for people to post and acquire digital information to and from public*

*digital displays, and to modify and annotate previously posted content to create publicly observable threads*". Therefore, they combined a public display with hand-held devices, which were used to augment, comment and annotate public content displayed on the large (public) screen.

Rekimoto (1998) proposed a multi-device approach for supporting informal meetings using digital whiteboards. In this scenario, tablet PCs were used as a data entry and tool palette for the (public) virtual whiteboard. As main functionalities for those coupled displays, text-entry (either by strokes or computer readable) and the retrieval of existing information were identified. In regard to text entry it is highlighted, that computer readable text has advantages over hand written text, as modifications to/search of existing textual content are possible. Concerning the information retrieval, Rekimoto points out that this can either target a private storage, public information services (such as the Word Wide Web) or data within the application itself. The coupled display is especially promising for search activity as performing a search directly on the public display is typically non-collaborative and therefore would disturb other participants' activities.

### 3. SCENARIO: A TABLETOP-BASED CREATIVITY SUPPORT SYSTEM

Basis of the work presented within this paper is a tabletop based application allowing for different, collaborative creativity techniques (Friess2010). Creativity techniques are typically based on a set of certain rules, activities and constraints aiming at providing a more structured process for creative problem solving as a form of guidance. These rules include heuristic principles such as forming associations, abstractions, analogies, combinations and variations. Our application is built on a (computer) model of those techniques which regards any creativity technique-based problem solving process as a sequence of two different types of process phases (Forster2009): In divergent phases, ideas for a given problem are sought, while in convergent phases, ideas from (preceding) divergent phases can be reviewed and evaluated.

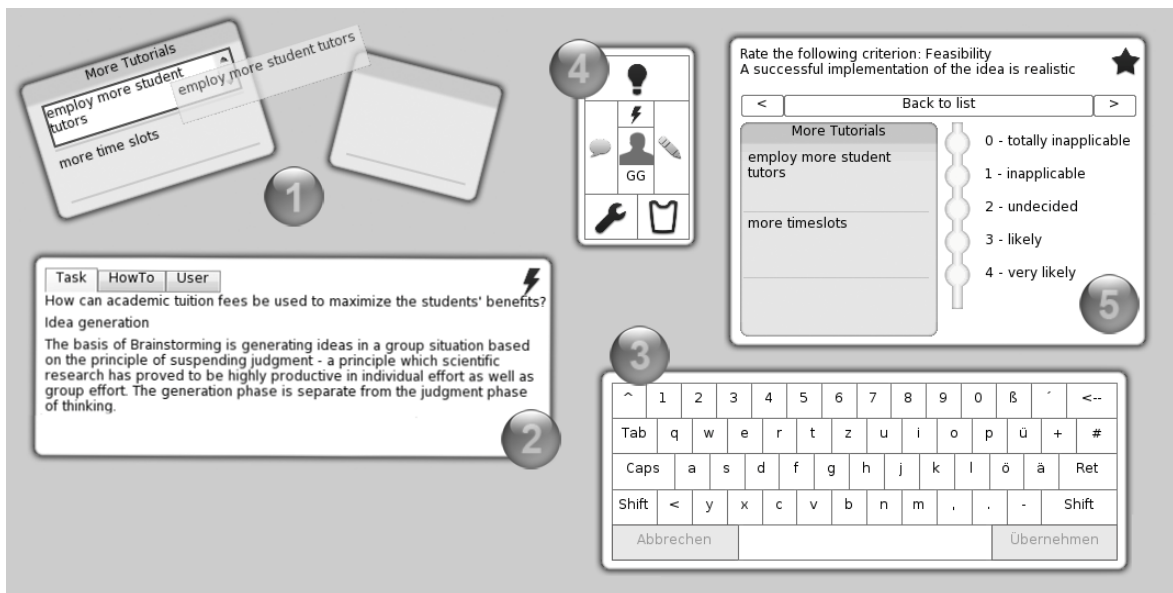


Figure 1. Widgets of the Underlying Tabletop based Creativity Support System

The user interface (UI) of the tabletop application is composed of so called widgets. Ideas form the product of any divergent phase and are represented as **(idea-)cards** (Figure 1, Item 1). As in divergent phases ideas get generated, such cards resemble modular bricks (so called "aspects") of textual content which can be manipulated (e.g. edited, reordered, moved or deleted). In addition to its aspects, each idea is assigned an individual title. Textual input is done via virtual (foldable) **on-screen keyboards** (Figure 1, Item 3), which, due to the multi-touch technology of the display, can be used concurrently. For convergent (idea evaluation) phases a separate **evaluation widget** (Figure 1, Item 5) is provided, which allows for assigning a rating to the ideas generated in a preceding divergent phase. We also included an **information widget** (Figure 1, Item 2)



which shows the current task as well as additional information given by the respective creativity technique (e.g. stimuli or constrains such as a time limit). Moreover, a **personal toolset** (Figure 1, item 4) allows each user to select an aspect or the idea title for editing by a drag and drop mechanism (by dragging from the light bulb icon to the respective area to edit). This way a user's virtual keyboard is connected logically to the area that is selected for modification and each contribution can be assigned to the corresponding user. An aspect which is marked for editing can also be moved to other ideas as presented in Figure 1, item 1. Moreover, the personal toolset allows a user to delete aspects or an idea as a whole (garbage can icon).

As the application also allows for participants to connect remotely to a co-located session by using a web-interface, a **chat widget** (not displayed) is provided for communication with such remote users. This widget can be switched in by pressing the speech bubble icon on the personal toolset. All widgets can be positioned and rotated arbitrarily, allowing the users e.g. to partition the workspace into territories (e.g. by rotating the ideas towards themselves) and to move freely around the table. For a more intuitive feeling we integrated a physics engine which allows each widget to be pushed over the table.

Based on the motivation presented in the introduction as well as by taking into account the related work presented in Section 2, we established a concept and a resulting implementation to improve this application by coupled displays. As a main question it had to be decided which actions / information should remain in the public space (= on the tabletop) and which be transferred to the private spaces (= the coupled displays). Two sub questions in this regard are how much functionality can be transferred without shifting the focus of the users away from the public display and how *invasive* the transferred actions should be. An interaction style focusing too much on the coupled displays may e.g. lead to a breakdown of the face-to-face communication as argued in the introduction. Allowing too invasive actions in private may result in disturbing/annoying other participants (e.g. anonymously deleting another user's idea could be problematic). Within the next section, the resulting application which was established by taking into account these considerations will be presented.

#### 4. THE COUPLED DISPLAY APPLICATION

In a conceptual view, we determined the following functionalities to be included: The coupled display should support **idea workspaces** (sets of ideas) being local (e.g. to work offline or completely in private) but also being connected to and synchronized with the public tabletop display. Within each workspace, it should be possible to select an idea and to get into a separate mode for editing, providing extended functionality to retrieve and create contents for an idea: On the one hand, the precise recognition of touch points allows for sketching while, on the other hand, the coupled display can also be used to include search functionality to e.g. search for images (either from local or remote storages) or (as another example) for ideas resulting from earlier sessions. In regard to the synchronized workspaces, we decided not to allow having all tabletop ideas directly on the coupled display per default. Instead, each idea must be specially selected in public (on the tabletop) to be shown (and edited) on the iPad. This way we intended to prohibit that participants isolate from the group and start working on the coupled displays only. Finally, several features which are only relevant for individual users can be outsourced to free up space on the tabletop display.

We decided to use the Apple iPad as coupled display device as we wanted to keep the same touch-based (natural and intuitive) style of human machine interaction which we already used for interacting with the tabletop application. We also see the iPad as a popular representative of the whole class of touch based tablet devices. Additionally, it provides a mature API to develop touch based applications. Also the larger display size made this category of devices a better choice (in regard to our requirements) than traditional smartphones. Using such devices is seen as intuitive to humans as it is similar to "*an oil painter [who] effectively uses a palette in his/her hand*" (Rekimoto1998, p.344).

To provide the groundwork for such a coupled display application, at first the tabletop application had to be extended for supporting images and sketches to be displayed as idea aspects. Moreover, we included a new button in the personal toolset allowing for *selecting* an idea to be observed on the coupled display. The communication with several coupled iPads was realized by a networking thread running in an endless loop and waiting for incoming or outgoing data to be processed. By default, a new server socket is opened while additionally, also the sockets of additional client connections need to be monitored. For communication, we chose to implement a custom TCP based binary protocol to avoid data overhead (e.g. imposed if we would

have used web service protocols). In this protocol, the general rules that apply to all packages are simple: A packet only has to supply a simple five byte header containing its length (including the size of the header; four bytes) and a numeric type identifier (one byte). As packets do not carry any further information about their content, this identifier is needed to enable the application to read the additional data. Basically, the implementation needs to provide some kind of lookup table that maps a packet's identifier to its structure. This way, the application can identify the start and end of a packet as well as validate its content to a certain level.

The iPad application itself is structured into several *IdeaSets* which logically group ideas as a horizontally scrollable list (see Figure 2). Most important is the **local *IdeaSet***, which allows for an offline (and private) creation, preparation, modification and collection of ideas. If an iPad gets coupled to the tabletop application, an *IdeaSet* of the **tabletop workspace** gets available. This set is by default empty and can be filled with ideas by selecting them for observation publicly. Publicly hereby means that a user has to select an idea visible to all participants on the tabletop. After this selection, he can edit the idea directly (and in private) on his iPad. Each committed change is then synchronized with the public display. An aspect which is currently edited in private also gets locked for other users. Searching for ideas is also realized by using *IdeaSets*. In this case an *IdeaSet* includes all ideas corresponding to an entered search criterion. Such a search can either target the local idea repository on the coupled display, the ideas of the current session (those on the tabletop surface) and ideas created in past sessions (the archive).



Figure 2. Coupled Display Application on the iPad (Text/Sketch Mode, Image Search, Rating)

Ideas (and their aspects) in the local and the tabletop *IdeaSet* can in general be edited, complemented and rearranged. For the local *IdeaSet* they can also be created, deleted and “pushed” to the tabletop display (by a dedicated push-button) to make them public. Ideas in *IdeaSets* resulting from a search query cannot be edited at all and only be copied. Editing an idea on the iPad works as follows: By tapping on an aspect a special view for modification becomes visible. There text can then be typed in (via the integrated onscreen keyboard) or it can be switched into a special mode for (touch-based) sketching (Figure 2, item 1). Images can only be added as new aspects by pressing the “pen”-shaped button (top right corner of an idea). Adding such an image is shown in Figure 2 (item 2). For the convergent (rating) phases, a rating bar is provided (Figure 2, item 3). In such phases, ideas cannot be edited anymore. Additionally convergent phases completely take place on the coupled displays, as no public space is needed.

On the top right of the user interface, several additional overlay-windows can be switched in. The information window shows the description of the current creativity technique and task while the chat window (the “speech bubble” icon) allows for communicating with remote users (or with co-located users in private). An online status field (bottom right) shows if the iPad is currently connected to a tabletop device.

## 5. EVALUATION

We evaluated the coupled display application involving 27 employees/students of our university. Those were partitioned into 7 groups resulting in 6 groups with 4 members and 1 group with 3 members. After a short introductory tutorial session, we chose a standard 20 minute long Brainstorming creativity technique for finding ideas to the problem “How could we improve the food supply at the campus”. In addition to the requirements of Brainstorming (Criticism is ruled out, Freewheeling is welcome, Quantity is wanted,

Combinations of ideas are sought) (Vangundy1988) and the 20 minute time limit no other constraints were imposed. In a next step, the group had to evaluate their ideas in regard to creativity and feasibility on a scale from 0 (worst) to 4 (best). Finally, we handed out a survey to each participant, asking about their opinion of aspects of the coupled display application and the group experience by using a Likert scale from -3 to +3 (“Strongly disagree” - ... - “Strongly agree”). In this regard one main aim was to find out about the participants’ perception of the invasiveness of certain actions (Q5), the isolation from the group (Q6) as well as the transitions between the different workspaces (Q7). Figure 3 shows a photo taken in the evaluation.



Figure 3. The Application in Action During the Experiment

Table 1. Selected Survey Results

	Question	Mean	stdev	Frequency (bold font = mode)						
				-3	-2	-1	0	+1	+2	+3
Q1	The user interface of the iPad was intuitive.	1.37	1.22	0	1	2	2	6	<b>13</b>	3
Q2	The additional features offered by the iPad improved the usability of the system.	1.74	1.35	0	2	0	1	6	<b>9</b>	<b>9</b>
Q3	The additional features offered by the iPad helped me to express my ideas better.	1.15	1.35	0	1	3	4	6	<b>9</b>	4
Q4	Working together with my group was fun.	1.63	1.42	1	0	2	1	4	<b>12</b>	7
Q5	I did not like other participants creating and modifying ideas anonymously.	-2.00	1.54	<b>16</b>	4	3	1	1	2	0
Q6	Participants were focused more on their iPad than on working with the group.	0.33	1.85	3	2	3	<b>6</b>	5	4	4
Q7	Switching between the iPad and the tabletop display was intuitive for me.	1.00	1.49	0	1	5	5	2	<b>10</b>	4

Throughout all sessions, a total of 91 ideas were generated, making an average of 13 ideas per group. The convergent (rating) phase resulted in an average creativity score of 2.5, while the feasibility score ranged at an average of 2.4.

Table 1 shows some selected results from the survey. A total of 81.5% of all participants rated the provided user interface as intuitive to use (Q1). Another 88.9% judged the additional features (e.g. the image/idea search, the sketching, etc.) as an improvement regarding the usability of the overall system (Q2). This is backed by 70% who agreed that those features are helpful to express ideas better (than just writing

down textual aspects by using the onscreen keyboard on the tabletop surface – Q3). Overall, 85.2% found the collaboration within their group (using the application) fun (Q4). Regarding the invasiveness of our approach, 85.2% stated that they did not have any problems with others creating and modifying ideas by using their coupled displays. However, the fact that the users were able to use the application in more ways than just as an extended editor lead to some users increasingly focusing on their coupled display and away from the group, as stated by 48,1% (Q6).

Summarized, the availability of coupled devices might always result in users isolating themselves from their group, consequently altering the character of the group's work to be less collaborative. However this negative effect is mitigated by extended, more parallel and more anonymous ways of input. Especially the image search was used frequently throughout all sessions resulting in many aspects containing images (cp. e.g. Figure 3). However, ideas containing images often were also extended by textual aspects which mainly described the idea behind the image. Switching between the different devices (the tabletop and the iPad - Q7) was stated as unproblematic by most participants, showing that our approach to use the same interaction style for both is promising. We also experienced that some users were using the infrastructure of the room (e.g. chairs) when they started working on their coupled display only (especially in the non-collaborative rating phases). Also, some users were working on the tabletop only in the beginning, switching over to their iPads when the public tabletop space got increasingly filled up with ideas.

Some participants stated that an idea that was prepared locally should be automatically observed when pushed to the tabletop as it is their "own" idea and they want to keep an eye on it. Most participants agreed that more invasive actions such as deleting ideas/aspects should be triggered only directly on the public (tabletop) display as it is the case in our application. Allowing deleting ideas in public only raises the inhibition threshold for performing such destructive actions on others' ideas as social norms indirectly steer co-located collaboration. Although the search for local / tabletop and archived ideas was used rarely, in other scenarios (e.g. when previous sessions were already conducted or when the participants remotely collected ideas previously, those features could be more important.

## 6. CONCLUSION

In the context of this article, a coupled display application extending a tabletop based CSS, which specifically targeted co-located collaborative creative work, was conceived. After presenting related work on multi-display environments in general, the application scenario in the form of a multi-touch table environment for supporting collaborative creativity techniques was introduced. Guided by the goal to mainly provide the users with more privacy and enhanced possibilities of input and information retrieval, the concept for and implementation of a coupled display application was presented. Following, an evaluation was conducted to investigate on the application and its influence on the (group) collaboration.

While this evaluation did show the general viability of the created multi display environment there clearly is still room left for further improvements and additional research on the influence of privacy and invasiveness. On the one hand the provided scenario (involving only a small number of students and using a very simple creativity technique) might not be really suited to investigate on e.g. the importance of contributing anonymously. In a company setting involving different hierarchy levels this might lead to more significant results. On the other hand, the private (offline) *IdeaSets* were hardly used in our evaluation. In different settings, e.g. combined with a preceding incubation phase, those might be used more actively and be more valuable. Regarding different creativity techniques under those aspects will be target of our future work.

## ACKNOWLEDGEMENT

The authors would like to thank InterFace AG (<http://www.interface-ag.com/>) for sponsoring the Apple iPads used in this research work.

## REFERENCES

- S. Carter, E. Churchill, L. Denoue, J. Helfman, and L. Nelson, 2004, Digital graffiti: public annotation of multimedia content. *CHI '04 extended abstracts on Human factors in computing systems*, pp. 1207-1210
- A.R. Dennis and B. Reinicke, 2004, Beta vs. vhs and the acceptance of electronic brainstorming technology. *MIS Quarterly*, vol. 28, issue 1, pp. 1–20.
- F. Echtler and S. Nestler and A. Dippon and G. Klinker, 2009, Supporting casual interactions between board games on public tabletop displays and mobile devices, *Personal and Ubiquitous Computing*, pp. 609–617
- Florian Forster and Michele Brocco. 2009, A unified process model for creativity-technique based problem solving processes. In *Marcus Specht Ulrike Cress, Vania Dimitrova, editor, Learning in the Synergy of Multiple Disciplines - 4th European Conference on Technology Enhanced Learning*, pages 615–620, Nice, France, Springer. EC-TEL 2009
- M.R. Friess, M. Kleinhans, F. Forster, F. Echtler, and G. Groh, 2010, A Tabletop Interface for Generic Creativity Techniques, *Proceedings of the IADIS Multi Conference on Computer Science and Information Systems (MCCSIS): Interfaces and Human Computer Interaction*, Freiburg, Germany, pp.203-210.
- S. Greenberg and M. Boyle, 1998., Moving between personal devices and public displays. *Workshop on Handheld CSCW*
- S. Greenberg, M. Boyle, and Laberge. 1999, PDAs and shared public displays: Making personal information public, and public information personal. *Personal and Ubiquitous Computing*, vol. 3, pp. 54-64
- O. Hilliges, L. Terrenghi, S. Boring, D. Kim, H. Richter, and A. Butz, 2007, Designing for collaborative creative problem solving, *C&C '07: Proceedings of the 6th ACM SIGCHI conference on Creativity & cognition*. New York, NY, USA, pp. 137–146.
- S. G. Isaksen, 1998, A Review of Brainstorming Research: Six Critical Issues for Inquiry, *Technical report, Creative Problem Solving Group*, Buffalo, New York
- N. Klügel, M. R. Frieß, G. Groh, and F. Echtler, 2011, An Approach to Collaborative Music Composition, *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*, Oslo, Norway
- C. Magerkurth, M. Memisoglu, Timo Engelke, and Norbert Streit, 2004, Towards the Next Generation of Tabletop Gaming Experiences. *Graphics Interface 2004 (GI'04)*, London (Ontario), Canada, May 17-19, pp. 73-80
- C. Magerkurth and T. Prante, 2001, "Metaplan" für die Westentasche - Mobile Computerunterstützung für Kreativitätssitzungen. *Tagungsband der GI-Fachtagung Mensch & Computer 2001 (MC'01)*, Bad Honnef bei Bonn, 5. - 8. März, Teubner Verlag
- J. Rekimoto, 1998, A Multiple Device Approach for Supporting Whiteboard-based Interactions. *Proceedings, CHI'98: Human Factors in Computing Systems*, pp. 344-351.
- G. Shoemaker, 2000, supporting private information on public displays. *Proceedings CHI EA '00: CHI '00 extended abstracts on Human factors in computing systems*, The Hague, The Netherlands, pp. 349 – 350
- D. Taylor, P. Berry, and C. Block, 1958, Does group participation when using brainstorming facilitate or inhibit creative thinking?, *Administrative Science Quarterly*, vol. 6, pp. 22–47
- A.B. Vangundy, 1988, "Techniques of Structured Problem-Solving", Springer Netherland, 1st edition

# ATTRACTIVENESS OF SOCIAL PLATFORMS: AN INVESTIGATION OF STUDI VZ

Waltraud Ilmberger, Martin Schrepp and Theo Held

SAP AG

Dietmar-Hopp-Allee 16, 69190 Walldorf

## ABSTRACT

We conducted an exploratory online study on StudiVZ, the most famous social platform for students in Germany, to find out which quality dimensions influence the perceived attractiveness of social software. The participants (n=722) were asked about duration, frequency and intensity of their StudiVZ usage and their preferred platform features. After a short screenshot walkthrough, the participants rated the user interface with the User Experience Questionnaire (Laugwitz, Held & Schrepp, 2008). The results indicate that features for establishing contacts were used predominantly. The hedonic quality dimension *stimulation* of the UEQ had most influence on general attractiveness of the user interface. This influence increases with the usage intensity of the social platform. The results indicate that investments in hedonic qualities pay off as they are very relevant, not only for a first impression but also for users that interact with the system a lot.

## KEYWORDS

Social networking software, attractiveness, user experience

## 1. INTRODUCTION

Web-based solutions for social communication, for example Facebook, MySpace or StudiVZ (a very popular German platform for students), become more and more popular.

Applications of this type are summarized under the term *social networking software*. They have different styles and address different target groups, but offer very similar functionalities (Dwyer, Hiltz & Passerini, 2007), like the ability to upload and share information, manage personal relationships, communicate, and to present the own person to other people. The aspect of creating and maintaining contacts is the main aspect of social networking software that distinguishes this type of products from blogs, wikis or social bookmarking solutions (Schmidt, 2006).

How do users judge the quality and attractiveness of the interface of social networking software? Currently there is not much information concerning this question available. A first analysis of relevant success factors can be found for example in (Preece, 2001). A recent study (Hart, Ridley, Taher, Sas & Dix, 2008) showed that the user interface of Facebook performed poorly with regards to traditional usability criteria evaluated by a heuristic evaluation using Nielsen's guidelines (Nielsen, 2011), but received very positive ratings by users concerning perceived ease of use. This could be explained by the fact that the more traditional usability criteria covered by the Nielsen principles are less important in this context than user experience factors, like for example fun, pleasure, identification (Hassenzahl, 2003), or surprise.

In this paper, we investigate the impact of quality aspects of the user interface of social networking software on the perceived attractiveness of the product. We are especially interested in how this impact changes with duration and intensity of use. We selected the social networking platform StudiVZ (<http://www.studivz.net/>) to investigate these questions in a large scale online study. StudiVZ is a quite popular social networking software in Germany. The main target group for this application are students of German universities.

## 2. RESEARCH QUESTIONS

The attractiveness of a user interface clearly depends on classical quality dimensions, like for example efficiency, effectiveness, consistency, or suitability for learning (for example, Molich & Nielsen, 1990). But in the last few years researchers in the field of human-computer interaction have increasingly discussed also the impact of aspects like joy of use (Hatscher, 2001), emotion (Norman, 2003), aesthetics (Kurosu & Kashimura, 1995; Tractinsky, 1997; Hassenzahl, 2002; Lindgaard & Dudek, 2003; Lavie & Tractinsky, 2004) or the influence of colors (Ilmberger, Schrepp & Held, 2008) on the subjective evaluation of a user interface. These aspects are often summarized under the term *hedonic quality*.

We use a research framework described in (Hassenzahl, 2002). This framework distinguishes between pragmatic quality dimensions, hedonic quality dimensions and the perceived attractiveness (attitude of the user towards the product, pure valence dimension) of a user interface. Pragmatic quality describes the task oriented quality aspects of the interface design, for example efficiency or effectiveness. Hedonic quality describes aspects that are not directly related to tasks the user performs with the product, for example novelty (is the design innovative and new or old fashioned) or stimulation (is it interesting or boring to work with the product). The research framework separates the pragmatic and the hedonic qualities from user interface attractiveness. It is assumed that the perceived attractiveness results from an averaging process of the perceived pragmatic and hedonic quality. Attractiveness can change if the subjective importance of these two quality factors changes. It is assumed that the perception of a user interface as pragmatic or hedonic is stable over time.

The main purpose of social networking software, like Facebook or StudiVZ, is to fulfill social needs concerning contacts and communication. It is typically not used to reach externally set goals or to work on concrete tasks. Thus, it is quite natural to assume that for this type of software hedonic quality dimensions have a higher impact on perceived attractiveness than pragmatic quality dimensions.

An interesting question is if and how the impact of hedonic and pragmatic quality dimensions on perceived attractiveness changes with the duration of usage. Sometimes product features create a stimulating user experience because they are perceived as new and surprising. If a user interacts with a product for a longer period of time, then this positive effect will naturally decrease. In contrast, problems concerning the interaction design become typically more visible the longer and more intense the product is used. In such a case we assume that the impact of hedonic quality dimensions decreases with increasing duration of usage while the opposite is true for pragmatic quality dimensions. Another factor that is investigated is the impact of the intensity of use on user expectations concerning hedonic and pragmatic quality. How intense a person uses social networking software obviously depends on the joy of use the person perceives during interaction. We investigate if users who use the product with high intensity have higher expectations concerning hedonic quality dimensions than users who use the product with lower intensity.

## 3. AN ONLINE STUDY FOCUSING ON *STUDIVZ*

The social networking platform StudiVZ was founded in 2005. Currently more than 10 million users (mainly students in German-speaking countries) take part in this network. The StudiVZ web page is accessed about 13.65 million times a month<sup>1</sup> – so it is one of the most frequently visited pages in German-speaking countries.

### 3.1 Participants

Participants were recruited by e-mails to large distribution lists and by postings containing a link to our study in newsgroups on StudiVZ. As a reward, participants had the chance to win a voucher for an online-shop. The landing page of the investigation was accessed 1339 times. The data sets of 722 participants (44% male, 56% female) were complete (drop-out-rate: 46.1%). Participants were between 15 and 51 years old (average:

---

<sup>1</sup> Source: AGOF internet facts 2010-II on [www.agof.de](http://www.agof.de) (January, 2010)

23.98, SD: 4.37). About two thirds (67.6 %) of the participants were students. Table 1 shows the information, participants provided regarding duration, frequency, and intensity of StudiVZ usage.

Table 1. Duration, frequency, and intensity of StudiVZ usage

Question	Categories	Quantity	Percentage
Usage duration	<3 months	36	2.9%
“How long have you been in StudiVZ”	3-12 months	157	21.7%
	>12 months	428	59.3%
	No answer	101	14.0%
Usage frequency	< 1 per week	77	10.7%
“How often do you log in at StudiVZ”	1 per week	56	7.8%
	> 1 per week	206	28.5%
	1 per day	170	23.5%
	> 1 per day	121	16.8%
	No answer	92	12.7%
Usage intensity (per week)	< 1 hour	301	41.7%
“How many hours do you spend at StudiVZ?”	1 – 2 hours	189	26.2%
	2 – 5 hours	95	13.2%
	> 5 hours	40	5.5%
	No answer	97	13.4%

## 3.2 Materials

We used the *User Experience Questionnaire (short UEQ)* (Laugwitz, Held & Schrepp, 2008) to investigate the participants’ impressions regarding the user interface of StudiVZ. This questionnaire consists of 26 items that have the form of a seven-stage semantic differential. An example item looks as follows:

attractive ①                      ②                      ③                      ④                      ⑤                      ⑥                      ⑦ unattractive

Those items measure the perception of a user interface regarding pragmatic and hedonic dimensions as well as a global dimension called *attractiveness*. Pragmatic dimensions are *perspicuity*, *efficiency*, and *dependability* and the hedonic dimensions are *stimulation* and *novelty*. *Attractiveness* is represented by six items (e.g., good – bad, attractive – unattractive), pragmatic and hedonic dimensions by four items each.

## 3.3 Procedure

The landing page of the investigation contained a welcome message and information regarding the thematic background, expected time required (10-15 minutes), and the reward for participation. The consecutive page contained a set of demographic questions and questions regarding the individual usage of StudiVZ (duration, frequency and intensity of usage). The participants were also asked which of the nine most prominent features of StudiVZ (for example, upload of photos) they use. The question was: *What are you usually doing in StudiVZ?*

To ensure that the participants’ ratings are in fact based on direct impressions from StudiVZ, typical user scenarios were presented as sequences of screen shots (see Figure 1 for an example). The scenarios were “sending a message”, “upload of a profile image”, “searching a member associated with a certain university”, and “putting a member on the ignore list”. Each scenario consisted of 4-8 screens. The scenarios were randomly assigned to the participants. After the scenario walkthrough, participants had to complete the UEQ.



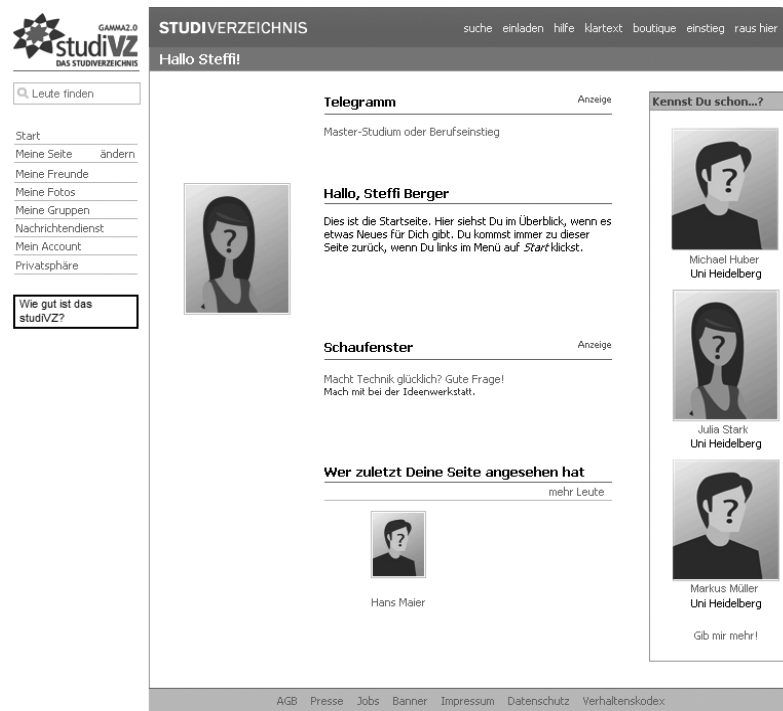


Figure 1. Screenshot presented in the online study

#### 4. RESULTS

Most of the participants use features for establishing and keeping contact (sending messages, searching for other students and getting birthday reminders). The second most used things are looking on other profiles, use discussion forums and upload images. The least used features were tagging photos, the fun feature “gruscheln” (greet and cuddle) and to find out who’s in your class. Figure 2 shows the frequency of the answers.

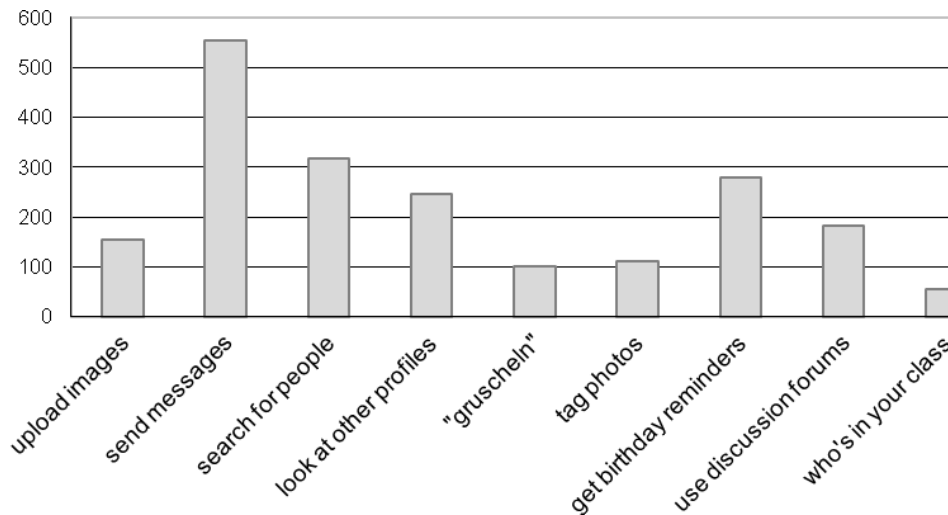


Figure 2. Frequencies for using the standard features of StudivZ

The means and standard deviations (in parentheses) of the UEQ dimensions for all 722 participants were: *attractiveness* 4.28 (1.08), *perspicuity* 5.35 (1.24), *efficiency* 4.34 (1.09), *dependability* 4.79 (.94), *novelty* 3.34 (1.06) and *stimulation* 3.64 (1.05). According to these results, the participants were rather satisfied with the pragmatic quality dimensions whereas the judgment of the hedonic quality dimensions was slightly negative (the range of values was from 1 to 7, therefore values around 4 would be a neutral judgment). Concerning the four used scenarios there were no significant differences with regard to the UEQ dimensions. Therefore, we do not differentiate between the scenarios in the following.

For testing the reliability of the single dimensions, Cronbach's Alpha was calculated per dimension. Following values were retrieved: *attractiveness* .90, *perspicuity* .85, *efficiency* .70, *dependability* .64, *novelty* .77 and *stimulation* .80. Besides the dimension *dependability* all dimensions have a satisfactory reliability.

Furthermore we found that *stimulation* strongly correlates with the three pragmatic quality dimensions (see Table 2).

Table 2. Intercorrelations between the dimensions of the User Experience Questionnaire

UEQ dimensions	Attractiveness	Stimulation	Novelty	Efficiency	Perspicuity	Dependability
Attractiveness	1					
Stimulation	.68**	1				
Novelty	.38**	.56**	1			
Efficiency	.69**	.47**	.22**	1		
Perspicuity	.64**	.35**	.06	.70**	1	
Dependability	.61**	.31**	-.04	.62**	.65**	1

\*\* significant with  $p < .01$

A linear regression of the five dimensions on the *attractiveness* was calculated to find out the influence of the particular hedonic and pragmatic dimensions on *attractiveness* of the user interface. The results of this analysis are shown in Table 3.

Table 3. Regression of the UEQ dimensions on attractiveness

UEQ dimensions	B	Std. Error	Beta
Attractiveness	-.68	.134	
Stimulation	.37	.03	.36***
Novelty	.14	.03	.14***
Efficiency	.19	.03	.19***
Perspicuity	.19	.03	.22***
Dependability	.28	.03	.25***

\*\*\* significant with  $p < .001$

All of the five scales have significant influence on the judgment of *attractiveness* ( $p < .05$ ), whereas *stimulation* has the most influence followed by *dependability* and *perspicuity*. *Novelty* had the least influence on *attractiveness*.

To investigate the influence of the usage duration, the participants were divided in three groups: *less than three months of usage* (36 participants), *between 3 and 12 months of usage* (157 participants) and *more than 12 months of usage* (428 participants). The means of the UEQ dimensions (see Table 4) didn't show any significant differences between the groups.

Table 4. Means and standard deviations (in parentheses) of the UEQ dimensions

Usage duration	Attractiveness	Stimulation	Novelty	Efficiency	Perspicuity	Dependability
<3 months (n = 36)	4.39 (1.35)	3.85 (1.50)	3.93 (1.33)	4.63 (1.26)	5.30 (1.49)	4.80 (1.13)
3-12 months (n = 157)	4.19 (1.21)	3.51 (1.14)	3.30 (1.15)	4.27 (1.12)	5.33 (1.30)	4.72 (.97)
>12 months (n = 428)	4.33 (1.00)	3.69 (.94)	3.33 (1.01)	4.36 (1.10)	5.40 (1.20)	4.82 (.90)

In the following step a regression analysis of the other dimensions on the dimension *attractiveness* was calculated per group. The results are shown in Table 5.

Table 5. Regression of the UEQ dimensions on the attractiveness separated to usage duration

UEQ dimensions	<3 months			3-12 months			>12 months		
	B	Std. Error	Beta	B	Std. Error	Beta	B	Std. Error	Beta
(constant)	-.05	.52		-1.21	.30		-.63	.19	
Stimulation	.39	.12	.44**	.32	.06	.30***	.35	.04	.33***
Novelty	-.01	.10	-.01	.25	.06	.24***	.13	.03	.13***
Efficiency	.44	.15	.41**	.18	.07	.16*	.15	.04	.16***
Perspicuity	.10	.14	.11	.15	.06	.17*	.24	.03	.29***
Dependability	.08	.18	.07	.40	.08	.32***	.26	.04	.24***

\* significant with  $p < .05$ , \*\* significant with  $p < .01$ , \*\*\* significant with  $p < .001$

Regarding the influence of the 5 dimensions on *attractiveness* there were strong differences between the three user groups. Within the group of new users only the dimensions *stimulation* and *efficiency* are relevant with regard to the *attractiveness* of the user interface. Both dimensions influence attractiveness equally. For the two other user groups all dimensions contribute significantly to the judgment of *attractiveness*. With medium usage duration (3-12 months) the dimensions *dependability* and *stimulation* are the most important ones. For long-term users (>12 months) *stimulation*, *perspicuity* and *dependability* are the most important dimensions. It should be considered that the three groups are strongly differing in size. An influence of the group size on the regression weights could be possible.

The importance of the dimensions *dependability* and *perspicuity* is rising with increasing duration of usage, which is not true for *efficiency*. To investigate the influence of the usage intensity two subsets were extracted from the 722 participants:

- *Frequent users*: participants that log on to StudiVZ more times a week or day and that use the platform more than 2 hours per week (134 participants).
- *Infrequent users*: participants that log on to StudiVZ once a week at most and use it less than 2 hours per week (132 participants).

Figure 3 shows the means of the UEQ dimensions for both groups.

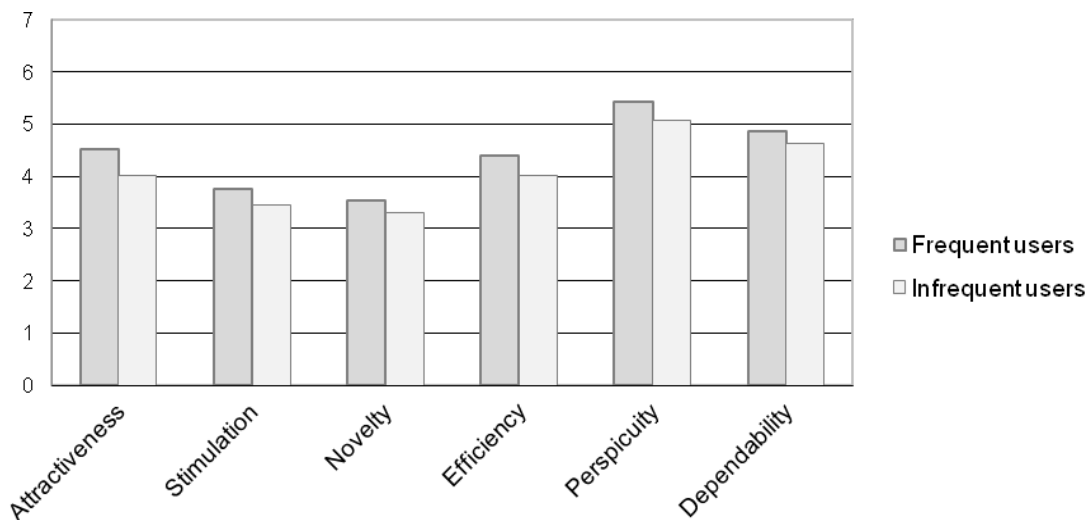


Figure 3. Scale means for frequent users and infrequent users

The mean values of the two groups only differ marginally (the differences for *attractiveness*, *efficiency*, *perspicuity* and *stimulation* are significant on the  $p < .05$  level). Frequent users evaluated all dimensions according to the expectations slightly better than infrequent users.<sup>2</sup>

<sup>2</sup> The observed mean values are (first value infrequent users, second value frequent users): attractiveness (4.01, 4.52), efficiency (4.00, 4.40), perspicuity (5.08, 5.42) and stimulation (3.46, 3.75)

A regression analysis showed that for both groups the *attractiveness* can be predicted by the other 5 dimensions (see also Table 6). For infrequent users *efficiency* has the most influence on the *attractiveness*. In contrast to this, the dimension *efficiency* is completely irrelevant for frequent users for the prediction of *attractiveness*. For infrequent users all dimensions have an almost equal influence on *attractiveness*. For frequent users *stimulation* and *perspicuity* are predominant. Novelty seems to be less important for both groups.

Table 6. Regression of the UEQ dimensions on attractiveness

UEQ dimensions	Infrequent users			Frequent users		
	B	Std. Error	Beta	B	Std. Error	Beta
(constant)	-.86	.31		-.54	.32	
Stimulation	.23	.06	.24***	.45	.07	.44***
Novelty	.20	.06	.22**	.12	.06	.13*
Efficiency	.28	.06	.30***	.06	.07	.06
Perspicuity	.18	.05	.24***	.29	.06	.33***
Dependability	.31	.07	.27***	.23	.08	.19**

\* significant with  $p < .05$ , \*\* significant with  $p < .01$ , \*\*\* significant with  $p < .001$

In comparison to infrequent users there is a stronger influence of hedonic quality dimensions on the *attractiveness* for frequent users. In contrast, pragmatic dimensions are more important for infrequent users than for frequent users.

## 5. DISCUSSION

Based on the example of the social network application StudiVZ we investigated, which of the quality aspects of the user interface of this type of software influence perceived attractiveness. Our results show that the hedonic dimension *stimulation* influences perceived attractiveness most strongly. This influence increases with intensity of usage. The result that importance of the dimensions *dependability* and *perspicuity* is rising with increasing duration of usage could be explained by a shift of importance within the pragmatic dimensions. Using a social platform, a user initially enters data to establish his profile. After this initial phase the relevance of data entering processes strongly decays. This could explain the decreasing importance of *efficiency* compared to the other two pragmatic quality dimensions with increasing duration of usage.

It has been interesting to observe that new users or users with low intensity of usage have a strong focus on *efficiency*, while *efficiency* seems to be irrelevant for long-time users or users with high intensity of usage when they are rating perceived attractiveness. This result can be explained with different *user modes* (Hassenzahl, Kekez & Burmester, 2002) of the user groups. According to this, users with low intensity of usage are supposed to work rather goal-oriented (*goal mode*) while users with high intensity of usage tend to interact with an interface less goal-oriented and more playful (*activity mode*). Users in activity mode rate the attractiveness of software more on basis of hedonic dimensions while users in goal mode focus on both, hedonic and pragmatic dimensions, when they are rating perceived attractiveness. The results indicate that investments in hedonic qualities pay off as they are very relevant for the user. This statement seems to be valid not only for a first impression but also for long-term users that interact a lot with the system.

## REFERENCES

- Laugwitz, B., Held, T. & Schrepp, M., 2008. Construction and evaluation of a user experience questionnaire. In: Holzinger, A. (ed.) USAB 2008. LNCS, vol. 5298, pp. 63-76. Springer, Heidelberg.
- Dwyer, C., Hiltz, S.R. & Passerini, K., 2007. Trust and privacy concerns with social networking sites: A comparison of Facebook and MySpace. In: Proceedings of the 13<sup>th</sup> Americas Conference on Information Systems, Keystone Colorado.

- Schmidt, J., 2006: Social Software: Onlinegestütztes Informations-, Identitäts- und Beziehungsmanagement. [Online based management of information, identity and relationships] *Forschungsjournal Neue Soziale Bewegungen*, No. 2.
- Preece, J., 2001. Sociability and usability: Twenty years of chatting online. *Behavior and Information Technology Journal*, 20(5), pp. 347-356.
- Hart, J., Ridley, C., Taher, F., Sas, C., Dix, A., 2008. Exploring the Facebook Experience: A New Approach to Usability. In: Proceedings of the 5<sup>th</sup> Nordic Conference on Human-Computer Interaction. ACM Press, New York.
- Nielsen, J., 2011. Ten Usability Heuristics. Available online under: [www.useit.com/papers/heuristic/heuristic\\_list.html](http://www.useit.com/papers/heuristic/heuristic_list.html) (last visited 12.01.2011).
- Hassenzahl, M., 2003. The thing and I: Understanding the relationship between user and product. In: Blythe, M., Overbeeke, C., Monk, A.F., Wright, P.C. (eds.) *Funology: From Usability to Enjoyment*. pp. 31-42. Kluwer Academic Publishers, Dordrecht.
- Molich, R., Nielsen, J., 1990. Improving a human-computer dialog: What Designers know about traditional interface design. *Communications of the ACM*, 33, pp. 338-348.
- Hatscher, M., 2001. Joy of use – Determinanten der Freude bei der Software-Nutzung. [Joy of use – Factors of fun by using software] In: Oberquelle, H., Oppermann, R., Krause, J. (eds.) *Mensch & Computer 2001: 1 Fachübergreifende Konferenz*. pp. 445-446. B.G. Teubner, Stuttgart.
- Norman, D., 2003. *Emotional Design: Why We Love (Or Hate) Everyday Things*. Boulder, Colorado: Basic Books.
- Kurosu, M., Kashimura, K., 1995. Apparent usability vs. inherent usability. In: *CHI'95 Conference Companion*, pp. 292 – 293.
- Tractinsky, N., 1997. Aesthetics and Apparent Usability: Empirical Assessing Cultural and Methodological Issues. In: *CHI'97 Electronic Publications*, Available online under <http://www.acm.org/sigchi/chi97/proceedings/paper/nt.htm> (last visited 10.2.2004).
- Hassenzahl, M., 2002. The effect of perceived hedonic quality on product appealingness. *International Journal of Human-Computer Interaction*, 13, 479-497.
- Lindgaard, G., Dudek, C., 2003. What is this evasive beast we call user satisfaction? *Interacting with Computers*, 15, pp. 429-452.
- Lavie, T., Tractinsky, N., 2004. Assessing dimensions of perceived visual aesthetics of web sites. *International Journal of Human-Computer Studies*, 60, pp. 269-298.
- Imberger, W., Schrepp, M., Held, T., 2008. *What kind of cognitive process causes the relationship between aesthetics and usability*. In: Holzinger, A. (ed.) *USAB 2008*. LNCS, vol. 5298, pp. 43-54. Springer, Heidelberg.
- Hassenzahl, M., Kekez, R., Burmester, M., 2002. The importance of a software's pragmatic quality depends on usage modes. In: Luczak, H., Cakir, A.E., Cakir, G. (eds.) *Proceedings of the 6th international conference on Work With Display Units (WWDU 2002)*. pp. 275-276. Berlin: ERGONOMIC Institut für Arbeits- und Sozialforschung.

# TOOL SUPPORT IN INTEGRATING TASK MODELING AND ACCESS CONTROL SPECIFICATION

Birgit Bomsdorf and Annika Paulheim

*Department of Applied Computer Science, University of Applied Sciences  
Marquardstr. 35, D-36039 Fulda, Germany*

## ABSTRACT

Security and access control, respectively, are often neglected in conceptual design of user interfaces. This paper presents a combination of HCI models and access control specifications to overcome this problem. The work aims at a systematic introduction of security concerns in model-based development of interactive systems. We concentrate so far on the early design steps, i.e. mainly on the specification of task and role hierarchies, and on conflicts that may result while the models are specified. Herewith, security concerns are considered at the stage of early design steps. The focus of this paper is on a tool supporting the integrative specifications of tasks and access control rules. It is shown how the tool assists in structuring the model taking into account model properties defined before. Incorporated functions check consistency violations and support to solve them. Additional support is realized by a view concept that assists the developer interactively in traversing the coherent model. In combination with filtering mechanisms and the possibilities of follow-up queries different views can be generated and altered. Both the automatic assistance in structuring a model and the dynamic view aim at the reduction of modeling complexity and at avoidance of modeling errors.

## KEYWORDS

Task modeling, Role modeling, Role task assignment, Tool support, Interactive model traversing, Access control.

## 1. INTRODUCTION

In model-based design, as performed in the field of HCI (see e.g. Calvary et al. 2003), it is well established to start with a model of the tasks users perform to reach their goals (task model). In addition, a model of the entities involved in task execution is created (domain and object model, respectively). Relations between task and object models describe how users may manipulate objects while performing the tasks. Parts of the user interface models (dialog and presentation model) are systematically derived from this information, but security policy is hardly based on it.

An access control model defines the permissions of users (e.g. human users, processes, computers) to access system resources (e.g. on an object, data base content, a file) (RBAC 2003). Role is taken as the central concept by which privileges to access objects and to perform operations on them are assigned to users. Generally, such models and security concerns, respectively, are often postponed until the end of the design cycle or until the implementation of a system (Díaz et al. 2008), (RBAC 2003). This holds true for the field in HCI as well, since security policy is often neglected in conceptual design of user interfaces. In Bomsdorf 2010 we have proposed a combination of HCI models and Role-Based Access Control (RBAC 2003) concentrating on the task layer as a first step to overcome this problem. Based on that work we modified the conflict algorithms and enhanced the view concept by interactive model traversing and a dynamic view concept, which was presented in this paper.

Our approach generally aims at the systematic introduction and tool support of security concerns in model-based development of interactive systems. Such a combination is not straightforward due to the differences in HCI modeling and access control: Although similar terms are used in HCI and access control their semantics differ. Particularly the concept role and privileges are used differently (Bomsdorf 2010). While role is the central concept in RBAC, HCI focuses primarily on tasks. Objects (also called task objects) are mostly of interest only if they are involved in task execution. Furthermore, task and object models are

formulated from the users' perspective in HCI whereas access control models represent a system-oriented view.

The basic concepts of our integrative approach comprise users, groups and roles, which are formalized according to RBAC (see Figure 1). The concepts task and object as defined in HCI are also included. Thus the *performs*-relation links the two "sides" together. Users are individual persons that are members of groups (see *belongs\_to* relation) whereby, because of *adopts* relations, they are enabled to perform tasks (*performs* relation) and hence to act on objects (*involves* relation), i.e. to invoke methods. Groups and roles are introduced to distinguish between roles that are based on the structure of an organization (groups in our approach) and roles based on the privilege to perform tasks, i.e. task-grouping roles (roles in our approach). Similar differentiation can be found in (Kang et al. 2001), (Osborn and Nyanchama 1999), (Romuald and Stéphane 2004) extending RBAC as well as the task-modeling approach TADEUS (Stary 2005).

Group, task and role models are structured hierarchically resulting in complex mutual dependencies. Guo (1999) has been showing in his work how this complexity may be handled in the context of access rights. Based on his proposal we extended our modeling approach (Bomsdorf 2007) to provide (semi-) automatic tool support in modeling the group, role, and task hierarchies and their relationships.

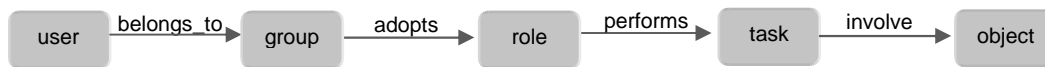


Figure 1. Overview of modeling concept

In RBAC privileges result mostly from role-resource assignments and are hierarchically organized. In our work access rights on objects result from assigning roles to tasks and tasks to objects. Comparable concepts can be found in extended models of access control, e.g. in (Kang et al. 2001) and (Zhang et al. 2000).

The task modeling tool CTTE (Paternò et al. 1998) enables to create separate task models per role defining all tasks that can be performed by that role. The role specification, however, does not support inheritance of privileges. This is supported in WSDM (Casteleyn and De Troyer 2001) but resulting role hierarchies are not formally integrated with task models.

UMLsec (Jürjens et al. 2008) aims at tool support for embedding security permissions in UML models. It enables, similar to conflict rules implemented by our tool (see below), to constrain security rules but does not consider privilege hierarchies. The approaches reported in (Díaz et al. 2008), (Garrigós 2003) and in (Romuald and Stéphane 2004) are comparable with our work. They aim at the integration of access control specifications with models known from Web-Engineering and Software Engineering, respectively. However, privileges are formulated by means of roles and system functionality, not considering the context of users performing tasks to reach goals.

This paper is structured as follows: Our integrated approach of task modeling and access control specification is present in the next sections with respect to tool support. We refer to it by the abbreviation RTME (Role-Task Model Extension) since the focus here is on tasks, roles and their mutual dependencies. In section 2 basic principles of task and role models are introduced showing how our tool enables specification and provides solution to conflicts in role-task modeling. In section 3 the interactive model traversing and the resulting dynamic view concept are introduced. This is followed by a final remark in section 4.

## 2. INTEGRATIVE MODELING WITH RTME

RTME supports the specification of group, role, and task hierarchies taking into account their interdependencies. Underlying functions check consistency violations and support to solve them. Hereby they contribute to the reduction of modeling complexity and to avoidance of modeling errors. The features are depicted in the following by means of examples. Since the main focus is on the relation between roles and tasks the basic concepts of role and task modeling are introduced as far as relevant for the subsequent presentation.

## 2.1 Interrelating Task and Role Hierarchies

Simplified task and role models of an online shop are used as a running example throughout this paper. Figure 2 shows the hierarchical structure of the exemplary task model. The task *online shopping* is a subtask of *online shop* and it is the superior task of *rate seller* and *buy product*, which is further refined. The task *sell online* is subdivided into two subtasks, whereas *administrate website* simply consists of a single task. Additional concepts are in use to specify the order of task execution, e.g. sequential or concurrent execution, and conditions constraining task performance, e.g. pre- and post-conditions (Baron and Scapin 2006), (Paternò et al. 1998), (Bomsdorf 2007). These concepts are not detailed here since in the subsequent considerations only the hierarchical task structure is of interest, i.e. they do not impact the privilege definitions and role conflicts.

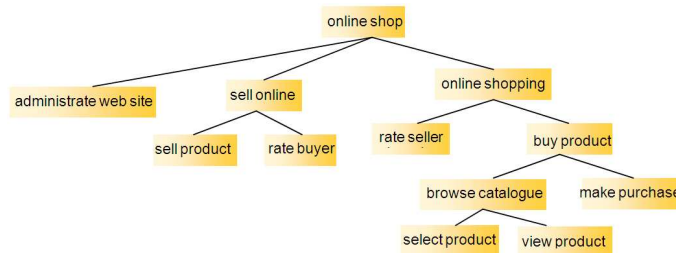


Figure 2. Example of a task model and role model

The concept *role* is used to define privileges. We denote the set of all privileges of a role  $r$  by  $privileges(r)$ . Roles are commonly organized hierarchically as well. In RTME a so-called poly-hierarchy is implemented that is represented by means of a graph. While a task hierarchy expresses composition relations, a role hierarchy describes inheritance of privileges. An edge from a role  $r_1$  to a role  $r_2$  indicates that the privileges assigned to role  $r_1$  are a proper subset of the privileges of  $r_2$ , i.e.  $privileges(r_1) \subset privileges(r_2)$ . For this kind of relation it is said that  $r_1$  is junior to  $r_2$  and that  $r_2$  is senior to  $r_1$  (RBAC 2001).

A role-task-mapping " $r$  performs  $t$ " (see Figure 1) means that a user who takes the role  $r$  possesses the privilege to perform  $t$ , and thus all its subtasks in the case  $t$  is further refined (Bomsdorf 2010). However, roles and tasks cannot be assigned arbitrarily to each other due to their hierarchical structures, i.e. it is not allowed to assign more tasks to a junior role than to one of its senior roles. This would result in an *assignment conflict* because in such a case  $privileges(senior\ role) \subset privileges(junior\ role)$  would hold true. This, however, is conflicting with the definition of the role hierarchy given above.

Modeling tools often implement a process in which developers define task and role models and specify the task-role assignments once the structure of the two is specified. This may result in unwanted assignments of privileges and assignment conflicts since the hierarchies and their interdependencies are often very complex. RTME follows the approach of automatically structuring the role model based on the task hierarchy specified so far to avoid such problems. How this works is shown by extending the online shop example.

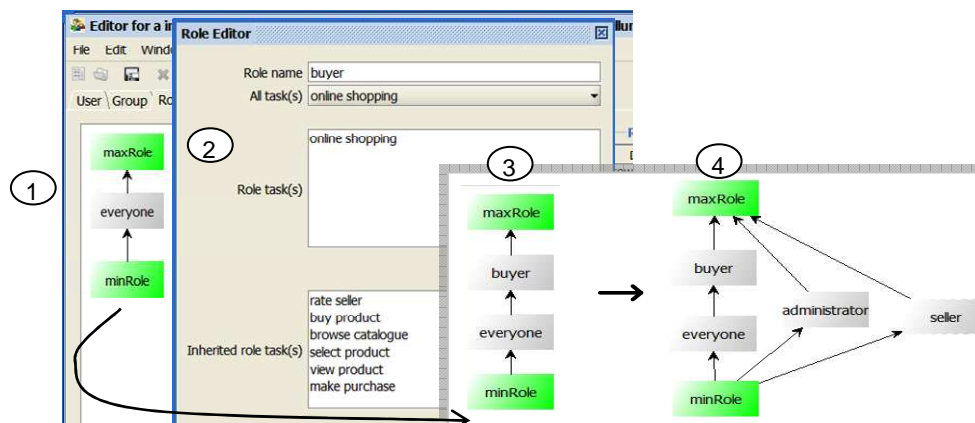


Figure 3. Example of role modeling



Let us assume we want to allow each visitor of the shop web site to browse the catalogue. However, only persons adopting a role *buyer* should be enabled to buy a product and to rate a seller. First of all we create a role *everyone* and assign to it the task *browse catalogue*. Figure 3 left hand (labeled with No. 1) shows the result of this editing step. Inserting the first role is done easily. It is positioned between *minRole* and *maxRole* as shown in the example with  $privileges(minRole) \subset privileges(everyone) \subset privileges(maxRole)$ .

Each graph possesses a minimal and a maximal role that are introduced for the purpose of computing a role model's hierarchy and taking care of conflicts. A detailed description is provided in (Guo 1999). Please note that the set of privileges of *minRole* is always empty while *maxRole* possesses all the time all privileges defined by the task model.

Now, see No. 2 in Figure 3, we create a role *buyer* and assign to it the task *online shopping* and hereby also all of its subtasks (see *Inherited role task(s)* in the role editor window). This step results in a model (see No. 3 in Figure 3) in which *buyer* is positioned into the role hierarchy according to the privileges added to the role as well as to the hierarchy existing so far. The task *browse catalogue* is part of *online shopping*, i.e.  $privileges(everyone) \subset privileges(buyer)$  holds true, based on which the position is determined. Similarly the roles *seller* and *administrator* are inserted to define privileges for *administrate web site* and *sell online*. The two roles are inserted separately into the role hierarchy since both *administrate web site* and *sell online* are sibling tasks of *online shopping* (see No. 4).

## 2.2 Conflicts

RTME checks for assignment conflicts each time the task or role model is modified. If a user of the role editor selects a task that would cause a conflict a warning is shown and the user is prompted to perform a correction.

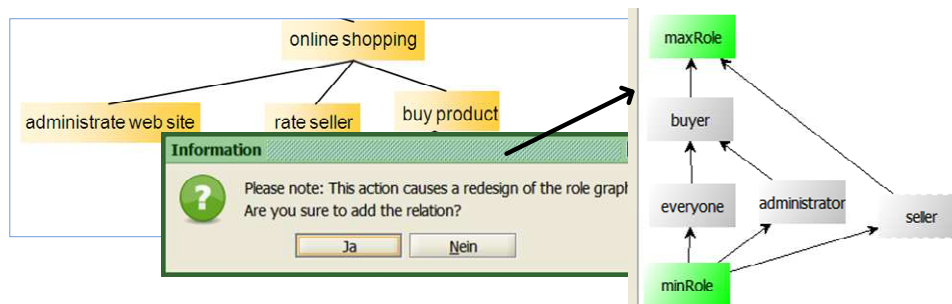


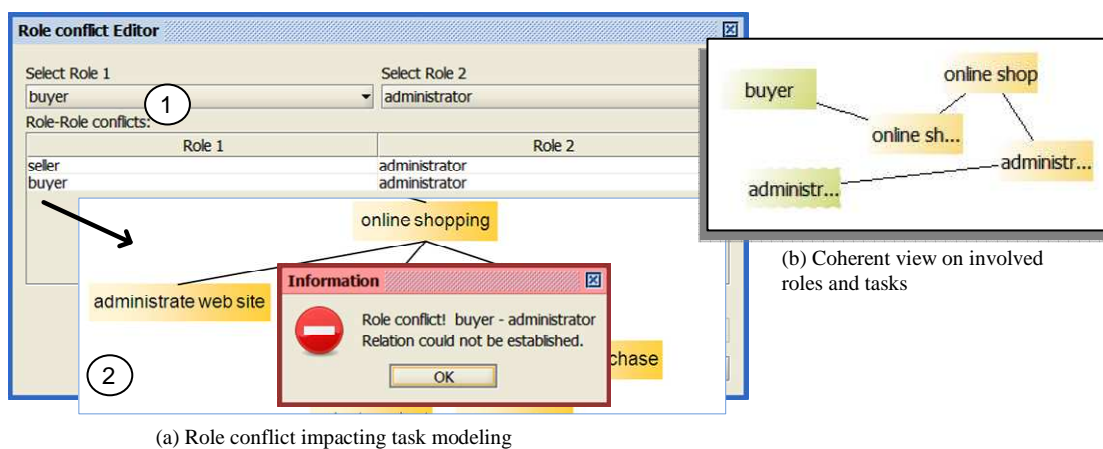
Figure 4. Role model re-configuration caused by task modeling

Similarly RTME evaluates editing steps while the task model is under construction. For example, we may change our task model so that *administrate web site* becomes a subtask of *online shopping*. The task editor, as shown in Figure 4 reacts to this by showing a dialog window. It notifies us that this modification will cause a redesign of the role model. In the scenario shown we agreed to the changes (by activating the *Ja*-Button meaning “yes”) and the role graph is re-structured. The role *administrator* becomes a junior role of *buyer*, i.e.  $privileges(administrate\ web\ site) \subset privileges(online\ shopping)$  holds true and thus a buyer is now allowed to perform all tasks of an administrator. However, this makes no sense in the online shopping example.

Users of RTME may define role conflicts to avoid such modeling cases. Generally, RTME enables to define explicit cases of conflicts, named *privilege conflicts* and *role conflicts* (Osborn 1999), which are checked during role and task modeling (Bomsdorf 2010). A *role conflict* is defined for two roles  $r1$  and  $r2$  if they should be mutually exclusive. In such a case a user must not take both roles. Basically, the two roles  $r1$  and  $r2$  must be independent from each other, i.e. there must be no role  $r$  (except *minRole*) that is junior of both at the same time. If so  $privileges(r) \subset privileges(r1)$  and  $privileges(r) \subset privileges(r2)$  would hold true. Hence a user who would take role  $r$  would possess privileges of both  $r1$  and  $r2$ , i.e.  $r1$  and  $r2$  are not mutually exclusive. In the case a role conflict is defined for  $r1$  and  $r2$ , in addition, *maxRole* must be the only senior role of  $r1$  and  $r2$ . If a role  $r$  would exist with  $privileges(r1) \subset privileges(r)$  and  $privileges(r2) \subset privileges(r)$ , a user taken  $r$  would possess privileges of both  $r1$  and  $r2$ .

A *privilege conflict* specifies two tasks  $t_1$  and  $t_2$  that must not be assigned to a role  $r$  at the same time. The role *maxRole* is an exception to this rule as it comprises all privileges existing in the model. However, this exception causes no problem since *maxRole* cannot be assigned to any user or group of users. Hierarchical dependencies, as in the case of role conflicts, have to be evaluated to guarantee fulfillment of specified privilege conflicts.

Obviously an explicitly defined conflict impacts role and task modeling. Figure 5 (see No.1) shows two role conflicts specified for the example given above: *seller* and *administrator* as well as *buyer* and *administrator*. Hence, common privileges of the roles *seller* and *administrator* as well as of the roles *buyer* and *administrator* are mutually exclusive. The attempt, for example, to define the task *administrate web site* as a subtask of *online shopping* results in an error now since  $privileges(administrate\ web\ site) \subset privileges(online\ shopping)$  must not hold true (see No.2). Please note that the task *online shopping* is assigned to the role *buyer* ( $online\ shopping \in privileges(buyer)$ ) while *administrate web site* is assigned to *administrator* ( $administrate\ web\ site \in privileges(administrator)$ ). Privilege conflict are similarly defined and handled by RTME.



(a) Role conflict impacting task modeling

(b) Coherent view on involved roles and tasks

Figure 5. Conflict examples

### 3. INTERACTIVE MODEL TRAVERSING

During editing, particularly if conflicts are detected, the user of RTME needs to know which elements and interdependencies have been specified so far. For example, in the scenario shown in Figure 5 (a) the developer gets a role conflict error message because he tries to establish *administrate web site* as a subtask of *online shopping*. He modifies the task model and gets information concerning the role model. The question is how the two models, more precisely how the task *administrate web site*, the task *online shopping*, the role *administrator* and the role *buyer* are interrelated so far causing such an error and how it may be solved. The developer may check within the task and the role editor, respectively, the properties of the involved model elements and their interrelations. For example, within the role editor developers may check relations to the task model by means of the role overview table (see Figure 6 left hand, No. 6). However, the property windows support only partial views but not a coherent overview including the assignments. The above Figure 5 (b) gives an exemplary coherent view showing that *administrator* is a role assigned to *administrate web site* and *buyer* is a role assigned to *online shopping*.

This is a simplified example but more complex views are needed for supporting the developers' modeling task. The complete model comprises the structured models of users, groups, roles, tasks and objects as well as their interrelations. Obviously, the model as a whole may be very complex, too complex to be shown completely with all elements and relations. However, in most cases only a restricted view on the model is required. For example, if the developer wants to know all tasks a user who has taken a role  $r$  is allowed to perform, only the tasks in  $privileges(r)$  are of interest. If the developer wants to know all roles a specific user is allowed to adopt, the related assignments *belongs\_to* and *adopts* are relevant. Similarly, in the example of

Figure 5 (a) the restricted view shown in (b) informs about the role-task dependencies with respect to the reported conflict. Furthermore, different questions come up from which follow-up questions may arise while the developer is investigating the model.

RTME provides a dynamic view concept to support developers in formulating different queries by which they may interactively traverse the coherent model. This feature is basically accessible through a query panel as shown in figure 6. It enables to formulate different queries according to which views are generated representing the results. Each query starts with a target type (see No. 1 in Figure 6) that defines the type of the elements to be contained in the result set of the query. This is followed by the type of the start element (No. 2) and the concrete instance of this type (No.3). The result can be visualized either as a *detailed view* or as *general view* (No.4). For example, the result of the query “Show all *tasks* assigned to *role buyer*” is shown in window A as well as in window B. Window A contains a detailed view representing all roles of the role hierarchy, all tasks of the task hierarchy and all assignments relevant in answering the question given by the query. Window B contains a general view showing only the task that is directly assigned to *buyer*.

Window C contains the general view that represents the result elements of the query “Show all *task objects* assigned to *user John McGill*”. In our example, although not introduced above, the user *John McGill* is member of the group *admin* and *registered user*. The current example contains no object assignments for the task of administrating the web site, thus the graph in window C contains no path from the group *admin* to any task object. Since *CD* is assigned to *online shopping* the whole path from user *John McGill* to the task object *CD* containing the involved group, role and task is represented. Window D shows the same result set but this time as a list. Generally, the developer can choose between a diagram and a list representation (see No. 5 in Figure 6).

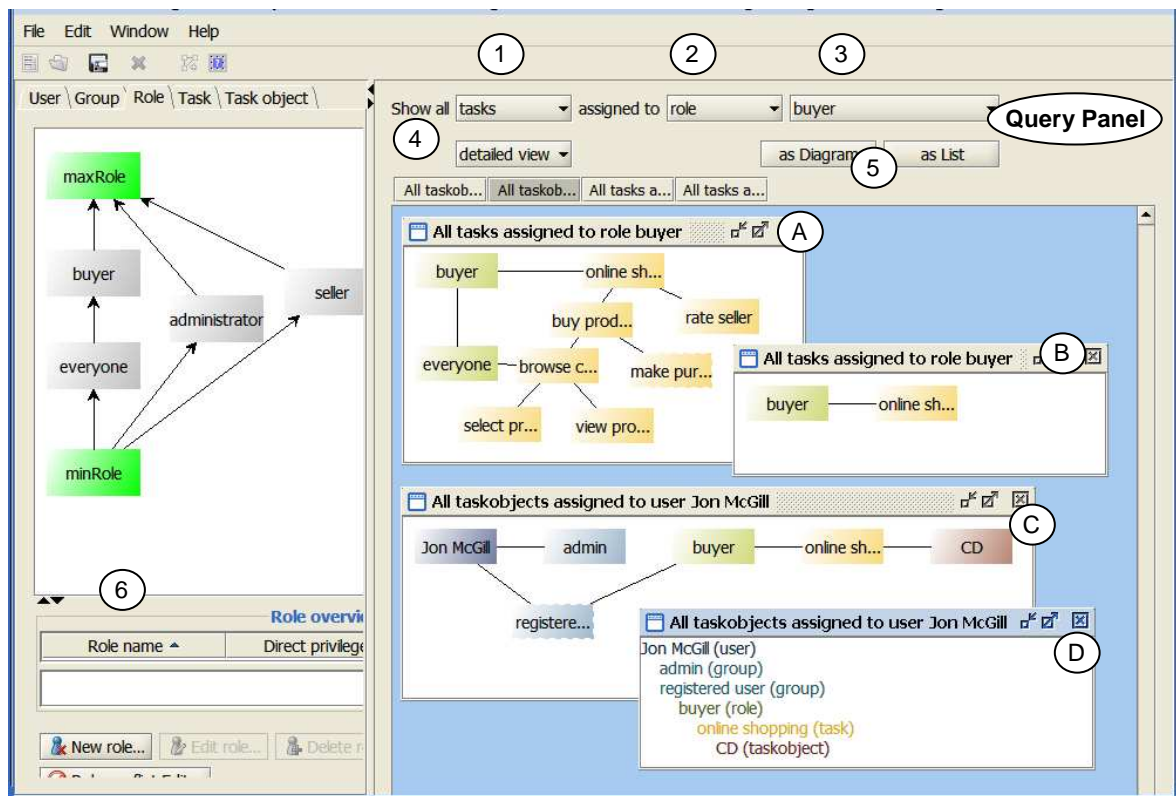


Figure 6. Overview and example of interactive model traversing

The developer may select any element representation, a node or a list entry, within a result window to manipulate the representation or to formulate a follow-up query. The respective features depend on the type of the selected element and are available through a context menu. It enables to show and hide single elements or all elements of a type. For example, if the context menu of a task is invoked its entries enable to hide the

task, its superior task, its subtasks, its directly assigned roles and task objects. If such elements are already hidden, further entries enable to show them.

Furthermore, a selected element can become the start of a follow-up query. Figure 6 shows results of such queries. The scenario starts in window A that contains the general view of the query “Show all roles assigned to task object CD”. Afterwards the task *online shopping* was selected and the context menu was invoked. Figure 8 No. 1 shows the subsequent selections of menu entries to formulate the follow-up query “Show all users assigned to this task”, i.e. to the task *online shopping*. The result set is to be presented by a new window as a diagram. This is shown in window B of Figure 7. Now we started in window B selecting *Jon McGill* and invoked the user context menu. This time (see No. 2, Figure 8) the follow up-query determines all roles of the selected user and the result is presented in the same window (see window C in Figure 7).

In the case the model is modified current views are automatically updated to show the modifications. For example, if the group *admin* is deleted its representation is removed in all current views. Additional representations of elements and links are removed or inserted according to the query a view is based on. In the example of deleting *admin* only the view in window C of Figure 7 would be affected, at which also the role *administrator* and the respective connections would be removed. If a view is altered only, e.g. by hiding selected nodes, the underlying model and all other current views remain unaffected. Modifying the model can be performed only by means of the model editors but not by the views.

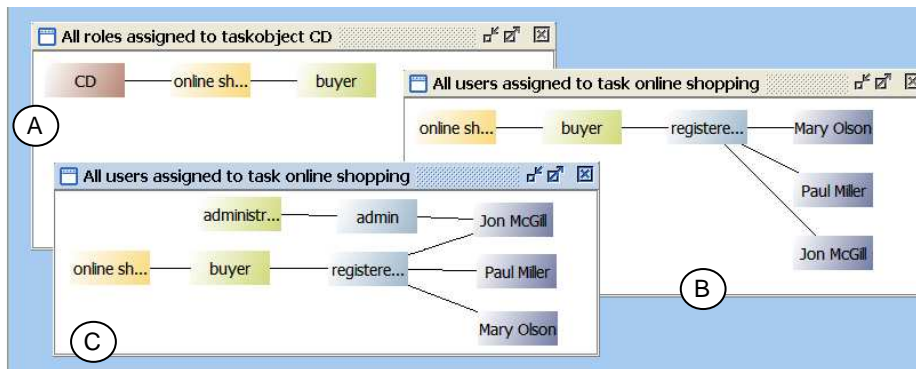


Figure 7. Examples of follow-up views

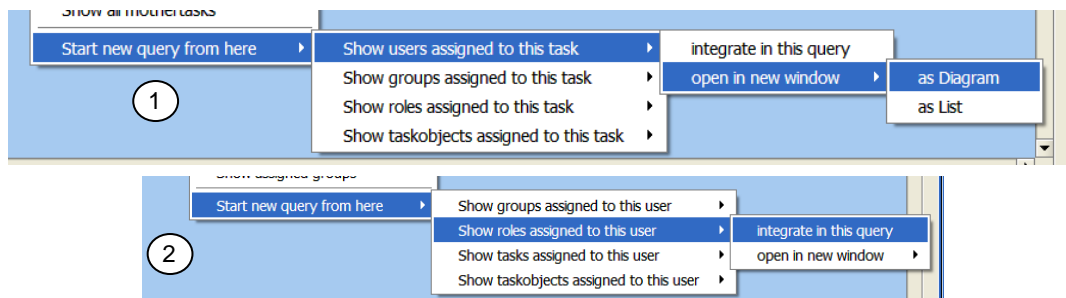


Figure 8. Formulating follow-up queries

## 4. CONCLUSION

Group models, as mentioned in the introduction, are also hierarchically structured. RTME handles the specification of group hierarchies and dependencies on role models in a very similar way as role-task modeling. All in all, our tool enables the specification of the groups a person belongs to, the roles a person may adopt, the tasks a person can perform, and the objects a person is allowed to access. The underlying theory combines our own work on task modeling (Bomsdorf 2007) with the work on conflict handling by (Guo 1999). Based on a first version of the tool RTME we modified the conflict algorithms and enhanced the view concept, which was presented in this paper.

RTME assists in structuring the model taking into account model properties defined so far. Each time a modeler inserts an additional assignment the resulting hierarchies are determined and the new assignment is inserted only if it causes no conflicts. This technique allows, for example, assigning roles and tasks to each other without performing required restructuring of the hierarchies. RTME evaluates instead the new graph models considering rules of conflicts that result from the hierarchical structures as well as those that are explicitly defined. This approach meets the fact that the model hierarchies result from the groupings of privileges.

Additional support is realized by the view concept that enables the developer to interactively traverse the coherent model. In combination with filtering mechanisms and the possibilities of the follow-up queries different views can be generated and altered. The interactive extracted views are presented in form of a list or by a diagram similarly to those used during editing. All the time the model is modified the presentation of the current views are automatically updated. More details on this approach are described in (Paulheim 2010).

The RTME editor is still in the state of a prototype. It possesses import and export functions (on the basis of XML) enabling the exchange of models with other tools. Currently our own task modeling approach and MAD (Baron and Scapin 2006) are supported. As further improvements of RTME we will work out the representations of the model elements and the views. Finally, we are aiming at its combination with our tool for model simulation to check effects of access rules during simulation task performance.

## REFERENCES

- Baron, M., Scapin, D.: K-MADe User Manual, <http://kmade.sourceforge.net>, (accessed 26.02.2011) (2006)
- Bomsdorf, B.: RTME: Extension of Role-Task Modeling for the Purpose of Access Control Specification, International Conference on Human-Centred Software Engineering, HCSE 2010, Springer, LNCS 6409, 149-157 (2010)
- Bomsdorf, B.: The WebTaskModel Approach to Web Process Modelling. In: TAMODIA Workshop, Springer, LNCS 4849, 240-253 (2007)
- Casteleyn, S.; De Troyer, O.: Structuring Web Sites Using Audience Class Hierarchies, In: DASWIS 2001 workshop (in conjunction with the ER 2001 conference), 198-211 (2001)
- Díaz, P., Aedo, I., Sanz, D., Malizia, A.: A model-driven approach for the visual specification of Role-Based Access Control policies in web systems, In: Visual Languages and Human-Centric Computing, VL/HCC 2008, 203 – 210 (2008)
- Garrigós, I., Gómez, J., Cachero, C.: Modelling Dynamic Personalization in Web Applications, LNCS, Vol. 2722, 3-27 (2003)
- Guo, Y.: User/Group Administration for RBAC. The University of Western Ontario (1999)
- Jürjens, J., Schreck, J., Yu, Y.: Automated Analysis of Permission-Based Security Using UMLsec, LNCS Vol. 4961, 292-295, (2008)
- Kang, M.H.; Parker, J.; Forscher, J.N.: Access Control Mechanisms for Inter-Organizational Workflow, In: Symposium on Access Control Models and Technologies, 66-74 (2001)
- Osborn, S.; Nyanchama, M.: The Role Graph Model and Conflict of Interest. The University of Western Ontario: s.n., (1999)
- Paulheim, A. Conception and Implementation of a Tool Supporting Interactive Model Traversing (in German: Konzeption und Implementierung eines Werkzeugs zur interaktiven Modelltraversierung), Bachelor Thesis, University of Applied Sciences, Fulda (2010)
- Paternò, F., Santoro, C. & Tahmassebi, S.: Formal Models for Cooperative Tasks: Concepts and an Application for En-Route Air Traffic Control. In: Design, Specification, and Verification of Interactive Systems, Springer-Verlag, Abingdon, 71-86 (1998)
- RBAC Standard, <http://csrc.nist.gov/rbac/rbac-std-ncits.pdf>, (accessed 26.01.2011) (2003)
- Romuald, T., Stéphane, C.: Integration of Access Control in Information Systems: From Role Engineering to Implementation, In: Informatica 30, 87–95 (2004)
- Sary, C.: Role-Adapted Access to Medical Data: Experiences with Model-Based Development, In: Universal Access in Health Telematics, 224-239 (2005)
- Stuart, J. & Penn, R.: TaskArchitect: taking the work out of task analysis. In: 3<sup>rd</sup> Annual Conference on Task Models and Diagrams, TAMODIA 2004, 145–154 (2004)
- Zhang, C., Hu, Y., Zahng, G: Task-Role Based Dual System Access Control Model, In: International Journal of Computer Science an Network Security, 6(7B), 211-215 (2006)

# MODEL-DRIVEN ENGINEERING OF BEHAVIORS FOR USER INTERFACES IN MULTIPLE CONTEXTS OF USE

Efrem Mbaki Luzayisu and Jean Vanderdonckt

*Université catholique de Louvain*

*Pole of Research on Information and Services Management and Engineering (PRISME)  
Louvain School of Management (LSM), Place des Doyens, 1 – B-1348 Louvain-la-Neuve (Belgium)*

## ABSTRACT

This paper describes a model-driven engineering approach for specifying, designing, and generating consistent behaviors in graphical user interfaces in multiple contexts of use, i.e. different users using different computing platforms in different physical environments. This methodological approach is structured according to the levels of abstraction of the Cameleon Reference Framework: task and domain, abstract user interface, concrete user interface, and final user interface. A behavior model captures the abstractions of the behavior in terms of abstract events and abstract behavior primitives in the same way a traditional presentation model may capture the abstraction of the visual components of a user interface. The behavior modeled at the abstract level is reified into a concrete user interface by model-to-model transformation. The concrete user interface leads to the final user interface running thanks to code by model-to-code generation.

## KEYWORDS

Behavior model, model-to-model transformation, model-to-code generation, user interface extensible markup language.

## 1. INTRODUCTION

For many years, the hardest part in conceptual modeling of User Interfaces (UIs) has been its dynamic part. All other aspects, such as presentation, help, tutorial, etc. have received considerable attention and results, especially in model-based approaches and model-driven engineering. We hereby refer to presentation as the static part of a UI such as the description of all windows, dialog boxes, widgets, and their associated properties. In contrast, we hereby refer behavior to as being the dynamic part of a UI such as the physical and temporal arrangement of widgets in their respective containers. The behavior has been also referred to as dialog, navigation, or feels (as opposed to look for presentation). Some typical behaviors are:

- When a language is selected in a list box, the rest of a dialog box is updated accordingly.
- When a particular value has been entered in an edit field, other edit fields are deactivated because they are no longer needed.
- When a validation button is pressed, the currently opened window is closed and another one is opened for pursuing the dialog.

The behavior received limited attention for many reasons: declarative languages that have been typically used for modeling presentation are hard to use for modeling behavior. Procedural languages could be used instead, but these induce a mixed-model-based approach that is complex to implement. Languages used for the final behavior are very diverse (markup or imperative), hold different levels of refinement (ranging from simple properties to sophisticated behaviors), are hard to abstract into one single level of abstraction (especially for different platforms), are hard to implement for model transformation. There is no consensus about what type of model should be used: some models exhibit a reasonable level of expressiveness, but prevent the designer from specifying advanced behaviors while other languages benefit from more expressiveness, but are more complex to handle, especially for non-trained designers. Which appropriate modeling approach is also open: taking the greatest common denominator across languages (with the risk of limited expressiveness) or more (with the risk of non-support), especially because many different implementations exist based on code templates and skeletons, deterministic algorithms, graph transformation, etc.

Finally, we are not aware of any existing approach that consistently applies model-driven engineering principles for UI behavior from the highest level (computing-independent model) to the lowest level (platform-specific model). Existing approaches only address some parts of some levels.

This paper is aimed at addressing these shortcomings by consistently applying principles of model-driven engineering starting with model-to-model transformation until model-to-code generation for multiple platforms. The rest of this paper is structured as follows: Section 2 summarizes the shortcomings of existing approaches. Section 3 introduces the conceptual modeling of behaviors used in this research. Section 4 presents an overview of the approach structured into steps that are supported by the software implemented. Section 5 explains the case study that we implemented to test methodology robustness. Section 6 concludes the paper and addresses some avenues.

## 2. RELATED WORK

Many different types of conceptual modeling and techniques of behaviors have been used over years [1-4, 7-21], some of them with some continuation, such as, but not limited to: Backus-Naur Form (BNF) grammars [9,12], state-transition diagrams in very different forms (e.g., dialog charts [1], dialog flows [3], abstract data views [7], dialog nets [8], windows transitions [14]), state charts [10] and its refinement for web applications [20], and-or graphs coming from Artificial Intelligence (e.g., function chaining graphs [19]), event-response languages, and Petri nets [2]. Comparing all these models in a sound way represents a significant contribution to what is yet to appear. Green [9] compared to a long time ago three dialog of them to conclude that some models share the same expressivity, but not the same complexity.

Cachero et al. examine how to model the navigation of a web application [4]. In [6], the context model is the basic model that could influence a behavior model at different steps of the UI development life cycle. So far, few attempts have been made to structure the conceptual modeling of behaviors in the same way as it has been done for presentation, the notable exception being applying StateWebCharts with Cascading style sheets [21] in order to factor out common parts of behaviors and to keep specific parts locally.

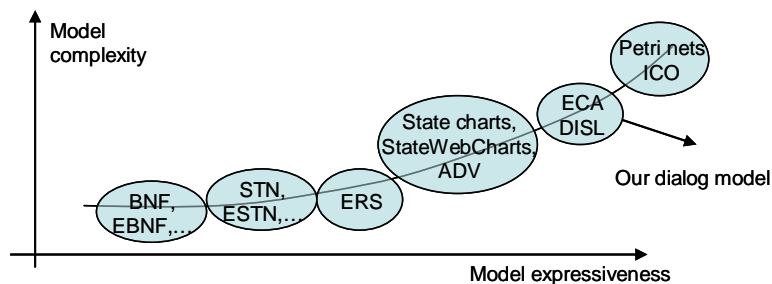


Figure 1. Model complexity as a function of their expressiveness.

Fig. 1 graphically depicts some behavior models in families of models. Each family exhibits a certain degree of model expressiveness (i.e., the capability of the model to express advanced enough behaviors), but at the price of a certain model complexity (i.e., the ease with which the dialog could be modeled in terms specified by the meta-model). At the leftmost part of Fig. 1 are located (E)BNF grammars since they are probably the least expressive behavior models ever. Then we can find respectively State Transitions Networks and their derivatives, then Event-Response Systems. Petri nets [3] are probably the most expressive models that can be used to model behaviors, but they are also the most complex to manipulate for inexperienced designers. Therefore, we believe that we could be less expressive and complex than Petri nets if Event-Condition-Action (ECA) systems are considered, such as in the DISL [18], UIML [11] and UsiXML (<http://www.usixml.org> - Fig. 2) User Interface Description Languages (UIDLs). In UIML, a behavior is defined as a set of condition-action rules that define what happens when a user interacts with an UI element, such as a button.

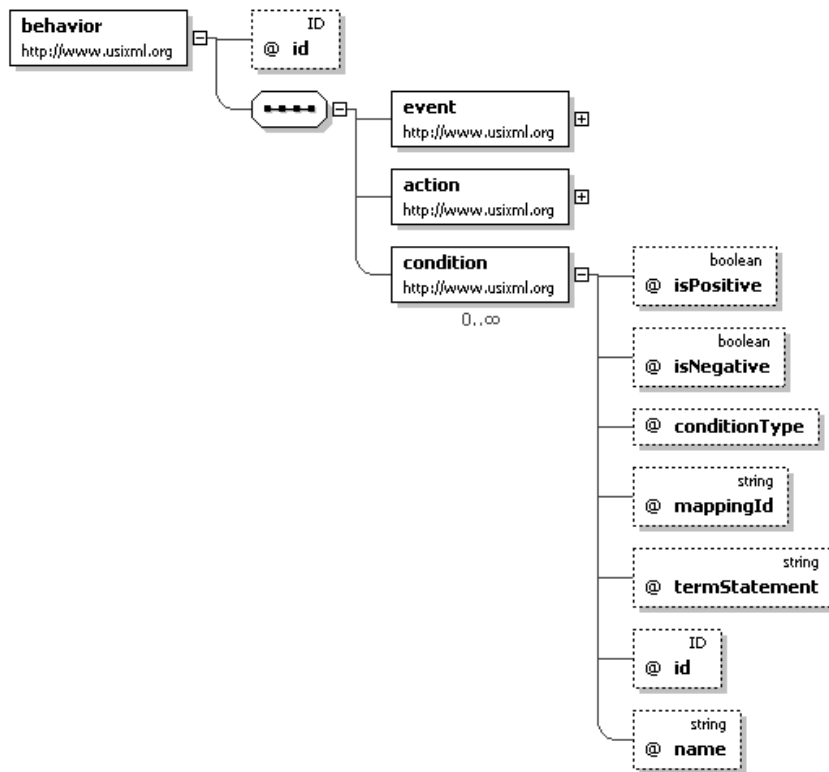


Figure 2. Modeling a behavior in UsiXML V1.6.8. Source: [http://www.usixml.org/documentation/usixml1.8.0/UsiXML.xsd.html\\_h-1518477259.html](http://www.usixml.org/documentation/usixml1.8.0/UsiXML.xsd.html_h-1518477259.html).

In UsiXML, a behavior (Fig. 2) is defined as a set of ECA rules, where: an event can be any UI event that is relevant to the level of abstraction where we are (it could be abstract or concrete).

- A condition can state any logical condition on a model, a model element, or a mapping between models.
- A condition is itself decomposed in term statements that can be reused, which are in turn decomposed into terms.
- An action can be any operation on widgets that are relevant to the level of abstraction where we are (abstract or concrete).

The UsiXML language describes the UI for multiple contexts of use such as character UIs, graphical UIs, auditory and vocal UIs, virtual reality, and multimodal UIs. As a language explicitly based on the Cameleon Reference Framework (CRF) [5], it adopts four development steps:

1. *Task & Concepts (T&C)*: describe the various users' tasks to be carried out and the domain-oriented concepts as they are required by these tasks to be performed.

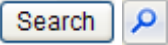

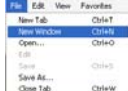

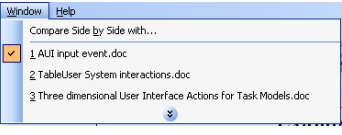


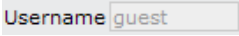

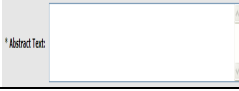

2. *Abstract UI (AUI)*: defines abstract containers (AC) and individual components (AIC) [15] two forms of Abstract Interaction Objects (AIO) by grouping subtasks according to various criteria (e.g., task model structural patterns, cognitive load analysis, semantic relationships identification), a navigation scheme between the container that selects the abstract individual component for each concept so that they are independent of any interaction modality, such as graphical, vocal, tactile, or haptic modalities. The AUI is said to be independent of any interaction modality.

3. *Concrete UI (CUI)*: concretizes an abstract UI for a given context of use into Concrete Interaction Objects (CIOs) so as to define widgets layout and interface navigation. It abstracts a final UI into a UI definition that is independent of any computing platform. While the AUI is independent of any interaction modality, a CUI assumes that a particular interaction modality has been chosen, but that this CUI remains independent of any computing platform.

4. *Final UI (FUI)*: is the operational UI i.e. any UI running on a particular computing platform either by interpretation (e.g., through a Web browser) or by execution (e.g., after compilation of code in an Integrated Development Environment (IDE)).



Table 1. Some mappings between abstract (AUI) and concrete (CUI) events

AUI input event	Example of CUI events		
	Graphical		Vocal
	GUI	Web	
Trigger an AIC with control facet	Click on a push button or icon 	Click on a push button  Open a window  Play a video file (start)	vocalInput and grammar and submit
Trigger an AIC with output facet		Scroll a web page without scrolling 	vocalPrompt or audio or vocalFeedback or vocalMenu with vocalMenuItems
Trigger an AIC with navigation facet	Select a view of an AC 	Select back or forward to navigate among pages  Select a view of an AC 	vocalNavigation or (vocalInput and grammar and connect)
Trigger an AIC with input facet		TextField single line in which the user can enter text.  PasswordField hides characters that user enters  TextArea in which the user can enter text 	vocalInput or record
Close an AC		Close or exit from a window 	Break or exit

### 3. MODEL-DRIVEN ENGINEERING OF BEHAVIORS

In order to apply traditional model-driven engineering techniques, we need to manipulate a behavior model that is expressive enough to accommodate advanced behaviors at different levels of granularity and different levels of abstraction, while allowing some structured design and development of corresponding behaviors. For this purpose, our conceptual modeling consists of expanding ECA rules towards *dialog scripting* (or *behavior scripting*) in a way that is independent of any platform (Fig. 3). For this purpose, we define five levels of granularity:

1. *Object-level behavior modeling*: this level models the behavior at the level of any particular object, such as a CIO or a AIO. In most cases, UI toolkits and Integrated Development Environments (IDEs) come with their own widget set with built-in, predefined behavior that can be only modified by overwriting the methods that define this behavior. Only low-level toolkits allow the developer to redefine an entirely new behavior for a particular widget, which is complex.

2. *Low-level container behavior modeling*: this level models the behavior at the level of any container of other objects that is a leaf node in the decomposition. Typically, this could be a terminal AC at the AUI level or a group box at the CUI level in case of a graphical interaction modality.

3. *Intermediary-level container behavior modeling*: this level models the behavior at the level of any non-terminal container of objects that is any container that is not a leaf node in the container decomposition. If the UI is graphical, this could be a dialog box or various tabs of a tabbed dialog box.

4. *Intra-application behavior modeling*: this level models the behavior at the level of top containers within a same interactive application such as a web application or a web site. It therefore regulates the navigation between the various containers of a same application. For instance, the Open-Close patterns mean that when a web page is closed, the next page in the transition is opened.

5. *Inter-applications behavior modeling*: since the action term of a ECA rule could be either a method call or an application execution, it is possible to specify a same behavior across several applications by calling an external program. Once the external program has been launched, the behavior that is internal to this program (within-application dialog) can be executed.

Now that have defined these five levels, we introduce the concepts used towards the conceptual modeling of behaviors that could be structured according to the five aforementioned levels of granularity. Fig. 3 depicts the global structuring of these concepts:

1. *Interactive Object*: interactive objects are particular objects used in the UI design and implementation. Strictly speaking, any way for a human being to “interface” with a machine is through a UI. The touch on its GPS or its microwave is a UI, the dial on its washing machine or the choice of television channels using a remote control are UI. Theoretically speaking, there exist interactive objects without physical representation. These objects are known as abstract. In the same way, there exists a classification of the objects according to their type. As all objects, interactive objects are characterized by their attributes (properties), behavior. Some objects have the capacity to react to the actions which they undergo by generating events. We can define a class as an abstraction of objects characteristics, including its attributes (fields or properties), its behaviors (methods or operators) and its events. In other terms, a class can be considered here as a model.

2. *Instance*: an instance is an individual object of a certain class. While a class is just the type definition, an actual usage of a class is called "instance". Each instance of a class can have different values for its attributes. At a given moment, the state of an instance is the set of its attributed values. By respecting the encapsulation i.e., the process of hiding all of the attributes of an object from any outside direct modification, object methods can be used to change an instance state.

3. *Toolkit*: in the modern graphical environments of design or development, Toolkits are represented as an objects box in which the developer can drag and drop items. In a UI Dialog Editor, the software that we developed in order to support this model-driven engineering of UI behaviors, the toolkit is characterized by its name, its level (e.g., a version), and a series of templates describing how this toolkit implements particular behaviors. Three values are accepted for the level: abstracted, concrete or final.

4. *Project*: any work session of UI Dialog Editor supposes the opening of a project. Indeed, a project can be viewed as a collection of interactive objects, or more precisely a collection of instances of interactive objects. To open and/or create a project, the user must first choose his toolkit.

5. *User Interface*: we define a user interface as the end result of a project. It represents a structured and harmonious set of interactive objects through which developer can implement interactive actions

6. *Mapping*: In the context of UI Dialog Editor, Mapping is a mechanism which makes it possible to define rules of transformation from a Toolkit towards another, possibly the same one. A given object can change with another or several others. In this second case, rules of transformations are necessary. They are written by means of the regular expressions applied to names of interactive objects. We can notice that the definition of a mapping is based primarily on two toolkits; a source and a destination. Its application requires presence of a user interface relating to the source Toolkit to lead to an interface in the destination Toolkit. It's possible to apply the same mapping into to several interfaces.

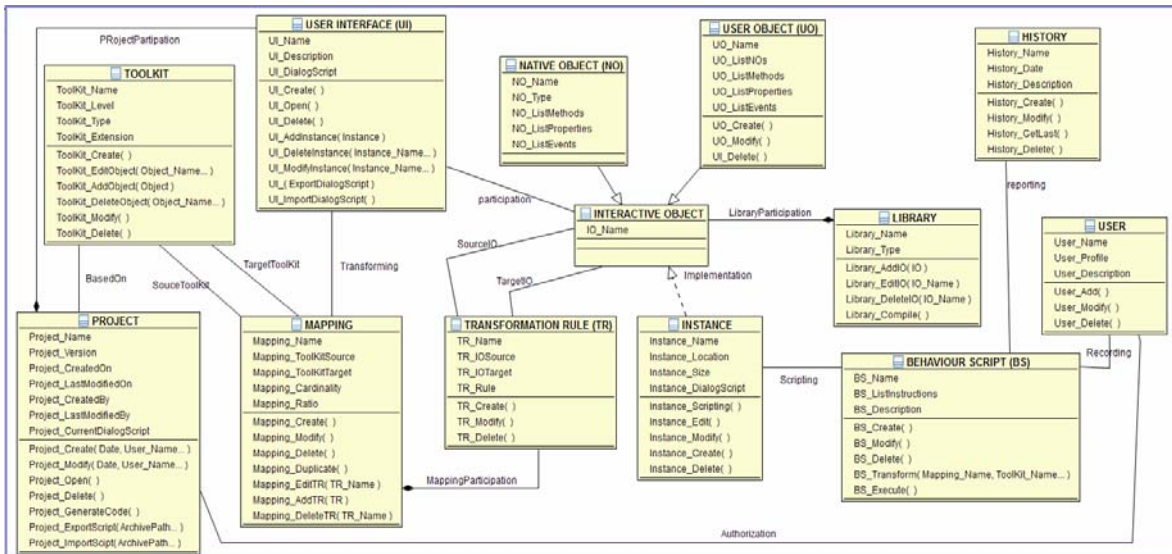


Figure 3. Conceptual modeling of behaviors for model-driven engineering.

7. *Behavior Script*: a script of dialogues is a sequential text furnished with the logic and conditional elements. It describes the desired actions according to a given interaction scenario. An action can be the change of an attribute value; the call of a mathematical function, the opening or the closing of a user interface, etc.

8. *History*: the history is the only means where we can ensure the constant traceability of scripts. It is possible to know which has modified which script of the dialogues and how. Thus, it will be possible to validate or cancel a given action.

We implemented a graphic dialog editor in which all these classes are exploited. This editor is presented in the next section.

## 4. DIALOG EDITOR IMPLEMENTATION

To support the above meta-model, we exploit MDA to build a graphic Dialog Editor in which Models are objects boxes, called Toolkit by abuse language. Models are organized in three levels (abstract, concrete and final according to CAMELEON architecture. Model elements are Interactive Objects (IO). In this context, a User Interface Project (UIP) is a sub-box in which all objects belong to a same given Model. In others terms, programming with Visual Basic 6 (VB6) and Visual Basic for Application (VBA), we implemented a dialog editor which

Is based primarily on the concept of objects Box. Developer has the choice between creating his own objects, using existing interactive objects or making both. It is important to determine which attributes, methods and events are necessary in dialog script.

1. Gives freedom concerning the level of specification. The user can choose to specify his project at the abstract, concrete or final level;

2. Generates the user interface at final level according to the platform toolkit. The resulting executable proposes architecture with three layers: The user interface which offers all interaction possibilities, the functional machine which contains semantic services and pilots database accesses and, the dialog controller which pilots exchanges between the two first layers;

3. Provides reification and concretization functionalities. Indeed, the user could, for example, use the same abstract specification to provide two or several different concrete specifications. In the same way, it could start from a concrete specification to lead to another concrete specification by skews of the mappings; and

4. Manages dialog scripts traceability. It is constantly possible to know who did what which day and at what time; and also, if necessary, it's possible to cancel recent modifications, or simply to revert to or resume with an old version of a given project

The Dialog Editor architecture is composed of four components presented in the figure 4 below.

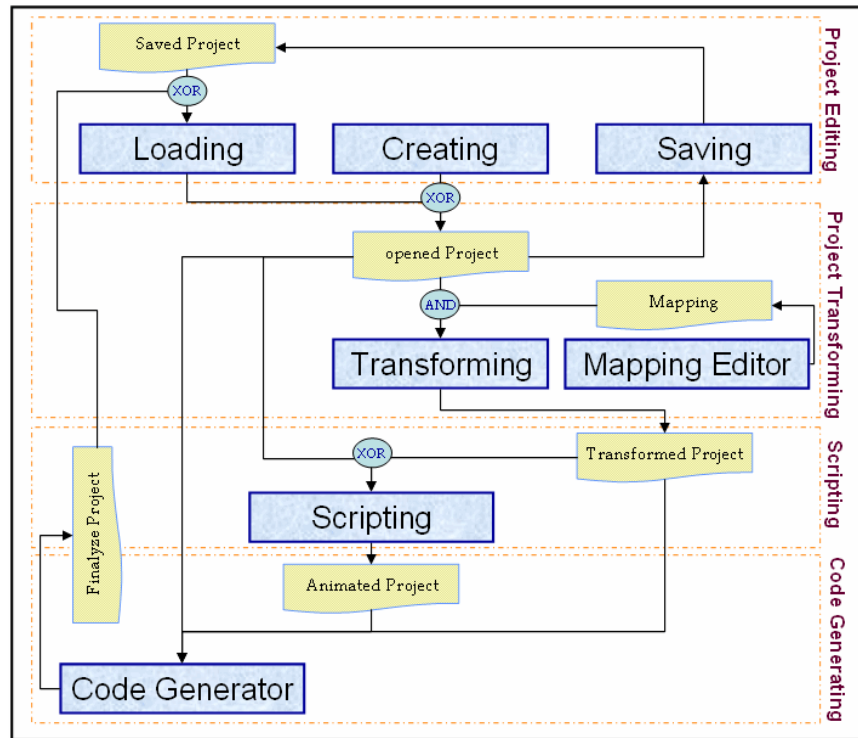


Figure 4. Dialog Editor Components

1. Project editing component proposes the functionalities concerning the creation of a new project. Firstly, it is necessary to choose the level of abstraction and the box with which to work. Then, by adding object after object, to build its project;

2. Project transforming component proposes mapping interfaces and functionalities of transforming a given project to another. It's possible to change specification level and/or specification toolkit.

3. Scripting component defines syntax and semantic of dialog language. In the objective to unify the specification of the interaction, we built a generic language which offers expressions for intra-object, intra-window, inter-windows, intra-process and inter-process dialog scripts. This generic language has a parser and a translator towards some computer programming languages.

4. Code generator integrate the algorithms which make it possible to transform a specification into a final project which is usable

We insist on the fact that the editor is developed primarily in Visual BASIC 6. Its interfaces are graphic. The internal data respect XML models. Looking in from the outside, the editor proposes four components or modules as shows in the dataflow diagram. The figure 5 below concerns project edition in BCHI Dialog Editor. It illustrates a project, named "106 – CTI Order AUI", where objects belong to "Abstract User Interface Model" Box. Each object is defined by its type, its name and its properties.

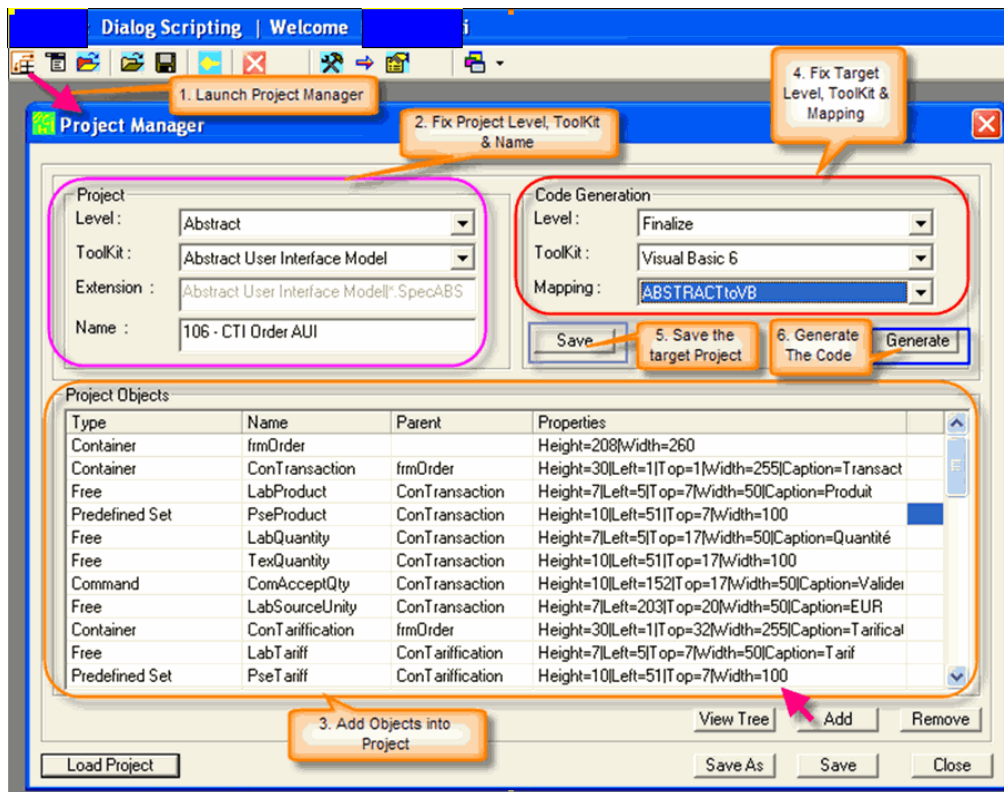
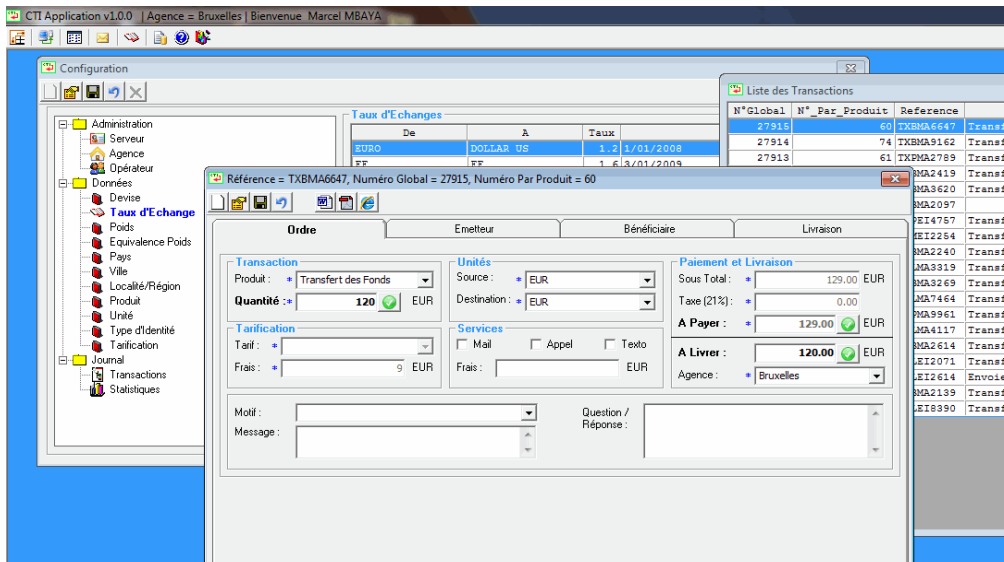


Figure 5. Project Interface

## 5. CASE STUDY

With our dialog editor, we developed a software aimed at covering the activities of a company which is specialized in the international transfer of money and import express worldwide services.



We needed an application based on a real case. But, we also needed an End-User oriented dialog application. For these two reasons, we asked the employers of CTI Company, whose head office is located at Liege in Belgium, to participate to the analysis and the validation of this software. Insofar as the employers of CTI Company speak only French, interfaces are in French in order to allow a better communication. In this context, commercial transaction is defined by: a Shipper (the customer which deposits the money or the object), a Sender (the person for which is intended the money or the object) and An Order (the details of transaction contents).

The recording of a new transaction requires other information which is saved beforehand in the database. Without being exhaustive, we can enumerate type of products, list of customers, list of currencies and their correspondences, tariffs, etc. Table 2 gives an outline temporary of the made tasks and, expressed as a percentage indicating the parts carried out manually and those generated automatically.

Table 2. Tasks distribution

Task	Timing	Manuel	Automatic
Interactive Objects library implementation	25%	100%	0%
Shipper, Sender and Order Management	65%	20%	80%
Reporting Management	10%	50%	50%

The interactive objects of the library which we have developed are useful both for the dialog editor and the CTI Application. To meet specific communications needs of company CTI, we integrated in this software, automatic functions of sending SMS and/or emails for transactions.

## 6. CONCLUSION

This paper introduces a particular method for supporting model-driven engineering of UI behaviors that are compliant with the Cameleon Reference Framework (CRF) [6]. For this purpose, a UI Dialog Editor has been implemented that: imports an AUI resulting from a generation from task and concepts, enriches it with dialog scripts that are abstract, apply model-to-model transformations in order to reify these scripts into concrete ones, and then apply model-to-code generation for three computing platforms. Three target contexts of user are supported: HTML for Applications (HTA) in a mobile context, Microsoft Visual Basic 6 and Microsoft Visual Basic for Applications (VBA) in a stationary context. Two operating systems are equally addressed: Microsoft Windows and Mac OS X. Five levels of behavior granularity are considered: object-level (behavior of a particular widget), low-level container (behavior of any group box), intermediary-level container (behavior at any non-terminal level of decomposition such as a dialog box or a web page), intra-application level (application behavior), and inter-application level (behavior across different interactive applications). Intra-container and inter-container behaviors are exemplified throughout a step-wide methodology that is supported by a dialogue editor, a model transformer, and a code generator, integrated into one single authoring environment.

The behaviors that can be modeled according to the approach presented in this paper can indeed lead to automated generation of their corresponding code. But these behaviors are for the moment straightforward and only handle general widget events such as activate, deactivate, etc. Very-fine grained behaviors affecting detailed properties of widgets are not supported.

## ACKNOWLEDGMENTS

The authors would like to acknowledge the support of the ITEA2-Call3-2008026 USI-XML (User Interface extensible Markup Language) European project and its support by Région Wallonne (DGO6). The authors would like to thank their wife and children for their great love and support.

## REFERENCES

- Ariav, G. and Calloway, L.-J. Designing conceptual models of dialog: A case for dialog charts, *SIGCHI Bulletin*, 20, 2 (1988) 23–27.
- Bastide, R. and Palanque, P. A Visual and Formal Glue Between Application and Interaction. *Journal of Visual Language and Computing*, 10, 5 (October 1999) 481–507.
- Book, M., Gruhn, V., and Richter, J. Fine-grained specification and control of data flows in web-based user interfaces. In *Proc. of ICWE'2007* (Como, 16-20 July 2007). LNCS, Vol. 4607, Springer-Verlag, Berlin, 2007, 167–181.
- Breiner, K., Maschino, O., Görlich, D., Meixner, G. Towards automatically interfacing application services integrated in a automated model based user interface generation process. In *Proc. of MDDAU'2009*.
- Cachero, C., Melia, S., Poels, G., and Calero, C. Towards improving the navigability of Web Applications: a model-driven approach. *European Journal of Informations Systems*, 16 (2007) 420–447.
- Calvary, G., Coutaz, J., Thevenin, D., Limbourg, Q., Bouillon, L., and Vanderdonck, J. A Unifying Reference Framework for Multi-Target User Interfaces. *Interacting with Computers*, 15,3 (2003) 289–308.
- Clerckx, T., Van den Bergh, J., and Coninx, K. Modeling Multi-Level Context Influence on the User Interface. In *Proc. of PERCOMW'2006*. IEEE, 2006, 57–61.
- Cowan, D. and Pereira de Lucena, C. Abstract Data Views: An Interface Specification Concept to Enhance Design for Reuse. *IEEE Trans. on Software Engineering*, 21,3 (1995) 229–243.
- Elwert, T. Continuous and Explicit Dialogue Modelling. In *Proc. of EA-CHI'96*.
- Green, M. A Survey of Three Dialogue Models. *ACM Transactions on Graphics*, 5, 3 (July 1986) 244–275.
- Harel, D. Statecharts: A visual formalism for complex systems. *Science of Computer Programming*, 8 (1987) 231-274.
- Helms, J., Schaefer, R., Luyten, K., Vermeulen, J., Abrams, M., Coyette, A., Vanderdonck, J. *Human-Centered Engineering with the User Interface Markup Language*. In “Human-Centered Software Engineering”, Chapter 7, HCI Series, Springer, London, 2009, 141–173.
- Jacob, R.J.K. A specification language for direct manipulation user interfaces. *ACM Transactions on Graphics*, 5, 4 (1986) 283–317.
- Lewis, J.R. IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use. *Int. J. of Human-Computer Interaction*, 7, 1 (1995) 57-78.
- Luyten, K., Clerckx, T., Coninx, K., and Vanderdonck, J. Derivation of a Dialog Model from a Task Model by Activity Chain Extraction. In *Proc. of DSV-IS'2003*. LNCS, Vol. 2844, Springer-Verlag, Berlin, 2003, 203–217.
- Mbaki, E., Vanderdonck, J., Guerrero, J., and Winckler, M. Multi-level Dialog Modeling in Highly Interactive Web Interfaces. In *Proc. of IWOST'2008*, CEUR Workshop Proc., Vol. 445, 2008, pp. 38-43.
- Menkhaus, G. and Fischmeister, S. Dialog Model Clustering for User Interface Adaptation. In *Proc. of ICWE'2003*. LNCS, Vol. 2722, Springer-Verlag, 2003, 194–203.
- Meixner, G., Görlich, D., Breiner, K., Hußmann, H., Pleuß, A., Sauer, S., Van den Bergh, J. Proc. of 4th Int. workshop on model driven development of advanced user interfaces. MDDAU'2009. In *Proc. of IUI 2009*, pp. 503-504.
- Meixner, G., Seissler, M., Nahler, M., Udit – A Graphical Editor For Task Models. In *Proc. of MDDAU'2009*.
- Pleuß, A. Modeling the User Interface of Multimedia Applications. In *Proc. of MoDELS 2005*, pp. 676-690.
- Pleuß, A. MML: A Language for Modeling Interactive Multimedia Applications. In *Proc. of ISM'2005*, pp. 465-473.
- Reichart, D., Dittmar, A., Forbrig, P., and Wurdel, M. Tool Support for Representing Task Models, Dialog Models and User-Interface Specifications. In *Proc. of DSV-IS'2008*. LNCS, Vol. 5136, Springer, Berlin, 2008, 92–95.
- Rückert, J. and Paech, B. The Guilet Dialog Model and Dialog Core for Graphical User Interfaces. In *Proc. of EIS'2008*. LNCS, Vol. 5247, Springer, 2008, 197–204.
- Schaefer, R., Bleul, S., and Müller, W. Dialog Modeling for Multiple Devices and Multiple Interaction Modalities. In *Proc. of TAMODIA'2006*. Lecture Notes in Computer Science, Vol. 4385, Springer-Verlag, Berlin, 2007, 39–53.
- Traetteberg, H. Dialog modelling with interactors and UML Statecharts. In *Proc. of DSV-IS'2003*. LNCS, Vol. 2844, Springer-Verlag, Berlin, 2003, 346–361.
- Vanderdonck, J., Limbourg, Q., Florins, M. Deriving the Navigational Structure of a User Interface. In *Proc. of Interact'2003*. IOS Press, Amsterdam, 2003, 455-462.
- Winckler, M. and Palanque, P. StateWebCharts: A formal description technique dedicated to navigation modelling of web applications. In *Proc. of DSV-IS'2003*. LNCS, Vol. 2844, Springer-Verlag, Berlin, 2003, 61-76.
- Winckler, M., Trindade, F., and Vanderdonck, J. Cascading Dialog Modeling with UsiXML. In *Proc. of DSV-IS'2008*. LNCS, Vol. 5136, Springer, Berlin, 2008, 121-135.
- W3C State Chart XML (SCXML), State Machine Notation for Control Abstraction, Working Draft, 16 May 2008. Accessible at <http://www.w3.org/TR/SCXML>.

# EMOTIONAL WEB-BASED DESIGN: THE CONCEPTS OF EMOTIONAL EXPERIENCE AND EMOTIONAL EXPRESSION

Zacharias Lekkas<sup>1</sup>, Nikos Tsianos<sup>1</sup>, Panagiotis Germanakos<sup>2,3</sup>, Constantinos Mourlas<sup>1</sup>  
and George Samaras<sup>2</sup>

<sup>1</sup>*Faculty of Communication and Media Studies, National & Kapodistrian University of Athens,  
5 Stadiou Str, GR 105-62, Athens, Hellas*

<sup>2</sup>*Computer Science Department, University of Cyprus, CY-1678 Nicosia, Cyprus*

<sup>3</sup>*Institute of Informatics and Telematics – CNR, 56124, Pisa, Italy*

## ABSTRACT

For many years people have been trying to measure differences between individuals. Over the course of time, a combination of developments in statistical know-how and the evolution of thought within psychology enabled the refinement of measures, and subsequently the assessment of more specific factors in the field of individual differences like different kinds of ability, affect and emotion. This knowledge has been used in many areas within psychology and at the same time the advancement of technology has enabled the development of web-based systems that measure specific factors relevant to specific situations. Our research focuses on the emotional mechanisms that drive human behaviour in general and how we can implement a set of rules to web design so that we can promote system adaptability on the very important field of human emotions which is at the same time extremely difficult to describe and define. In this paper we introduce our model of emotion regulation and we present our first experimental results that concern the concepts of emotional experience and emotional expression and their effect on decision making and problem solving styles. Furthermore, we present the implications that these theoretical and empirical representations can have in web applications and design.

## KEYWORDS

Emotion regulation, emotional experience, emotional expression, decision making, problem solving

## 1. INTRODUCTION

Since 1994, the Internet has emerged as a fundamental information and communication medium that has generated extensive enthusiasm. The Internet has been adopted by the mass market much quicker than any other technology over the past century and is currently providing an electronic connection between progressive businesses and millions of customers and potential customers whose age, education, occupation, interest, and income demographics are excellent for sales. The explosive growth in the size and use of the WWW as well as the complicated nature of most Web structures result in orientation difficulties, as users often lose sight of the goal of their inquiry, look for stimulating rather than informative material, or even use the navigational features unwisely. As the e-Services sector is rapidly evolving, the need for such Web structures that satisfy the heterogeneous needs of its users is becoming more and more evident.

To alleviate such navigational difficulties, researchers have put huge amounts of effort to identify the peculiarities of each user group. Their goal is to analyze and design methodologies and systems that could deliver up-to-date adaptive and personalized information, with regards to products or services (Mulvenna et al, 2002). Further consideration and analysis of parameters and contexts such as users' cognitive and mental capabilities, socio-psychological factors, emotional states and attention grabbing strategies should be extensively investigated. All these characteristics could affect the apt collection of users' customization

---

This work is co-funded by the EU project CONET (INFSO-ICT-224053) and by the project smarTag (within the framework of University of Cyprus internal funded projects).



requirements and along with the ‘traditional’ user characteristics (i.e. name, age, education, experience, profession etc.) constitute a comprehensive user profile that serves as the ground element of most of these systems offering in return the best adaptive environments to their preferences and demands.

Adaptivity is a particular functionality that alleviates navigational difficulties by distinguishing between interactions of different users within the information space (Eklund & Sinclair, 2000). The user population is not homogeneous, nor should be treated as such. To be able to deliver quality knowledge, systems should be tailored to the needs of individual users providing them personalized and adapted information based on their perceptions, reactions, and demands.

Web-based information systems are increasingly being used for learning and training applications. Computers are becoming better and more sophisticated every day. They can already perceive information related to user needs, preferences and characteristics (Cingil et al, 2000; Kim, 2002). One possible implementation of a Web-based system’s interface that can appraise human characteristics is through the use of a series of online tests and questionnaires that can assess the psychological abilities and properties of the user (Picard, 1997).

The concepts of emotion and affect underpin psychology’s attempt to identify the unique character of individuals. The terms describe properties of behavior which concern the individual’s typical ways of coping with life events (Lewis & Haviland-Jones, 2004). Norman (2004) argued that in order for media to communicate better with people they need to be able to understand our emotion and in order to do that they need to have emotions as well. An in-depth model that grasps the complexity of these underlying concepts is the first purpose of our research. Instead of selecting one area of implementation we are trying to combine various levels of analyses and form a typology that will help us circle effectively the cognitive and affective mechanisms of the brain. In order to apply a purely psychological construct to a digital platform we adjust the various theories concerning cognition and emotion having in mind to make our model flexible and applicable to users’ profiles, needs and preferences. In order to manipulate the emotional parameters according to user characteristics, our research has to go through the stage of extracting quantified elements that represent deeper psychological and affective abilities. The latter cannot be directly used in a web environment, but a numerical equivalent can define a user characteristic. We developed a theory and a corresponding battery of questionnaires for the concept of Emotion Regulation. Our psychological model has two base elements: The experiential level which is the actual emotional experience and emotional expression of the individual (the capacity of a human being to sense, experience and express specific emotional situations) and the rational level which is the multiple ways with which the individual recognizes and manages emotions. An effort to construct a model that predicts the role of specific emotions is beyond the scope of our research, due to the complexity and the numerous confounding variables that would make such an attempt rather impossible. We focus on emotion regulation as an emotional mechanism and not on a number of basic emotions because experiential emotion regulation can provide some indirect measurement of general emotional mechanisms since it manages a number of emotional factors like anxiety, boredom effects and frustration. Our model would be problematic without a regulatory mechanism of emotion. For this reason we included also the rational level of Emotion Regulation that is comprised from terms like emotion recognition, emotional management and emotional motivation. Emotion regulation is the way in which an individual perceives and controls his emotions. Individuals attempt to influence which emotions they have, when they have them and how they experience and express them (Bechara et al, 2000). By combining the affective state of the individual with his regulatory mechanism we can reach into a conclusion of how emotions influence his performance and the outcome of his behavior.

## **2. THE MODEL OF EMOTION REGULATION**

Theorists from a variety of orientations tend to agree in two emotional processing systems. There is considerable conceptual overlap in their formulations:

- A schematic, associative and implicit system that has connections with bodily response systems. This mode involves fast and automatic processes such as priming and spreading activation. It often involves large numbers of memories in parallel. It is not wholly dependent on verbal information – visual, kinaesthetic or other cues could provide the basis for priming or activating an emotional memory.

- An abstract propositional ‘rational’ system that is analytical, reflective, logical and relies on high level executive functions. It is primarily based on verbally accessible semantic information.

Individuals can utilize these two systems to process information. The first system relies on experience and intuition. In particular, individuals consider issues intuitively and effortlessly. Rather than reflect upon the various considerations in sequence, individuals form a global impression of issues. In addition, rather than apply logical rules or symbolic codes, such as words or numbers, individuals consider vivid representations of objects or events. These representations are filled with the emotions, details, features, and sensations that correspond to the objects or events. Finally, learning is equated to ascertain associations from direct experiences.

The second system, in contrast, relies on logic and rationality. In particular, individuals analyze issues with effort, logic, and deliberation rather than rely on intuition. To decide upon issues, they rely on logical rules and symbolic codes. The context (details, features, and emotions) that correspond to objects or events are disregarded. To facilitate learning in this system, individuals learn the rules of reasoning that are promulgated in society.

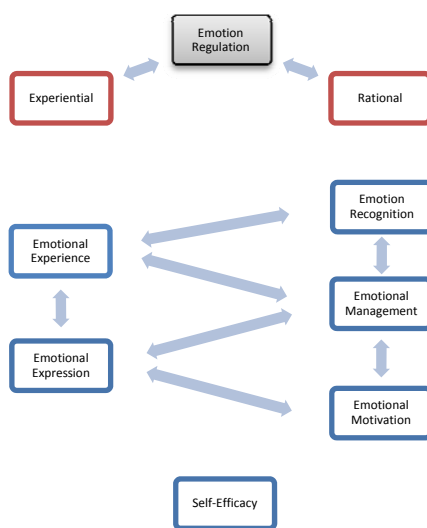


Figure 1. The Emotion Regulation Model

Recent neuroscientific findings are consistent with these multi-level conceptualisations. Le Doux (1998) has reviewed evidence suggesting that emotion networks have direct anatomical connections to both the neocortex and the amygdala. Events that are highly emotional are likely to be registered at both subcortical and cortical levels. The subcortical route is shorter and rapid whereas the cortical route is longer and slower. In the subcortical route sensory information goes from the thalamus directly to the amygdala. In the cortical route information is sent from the thalamus to both the cortex and hippocampus and is then projected to the amygdala. As noted by Samilov & Goldfried, (2000) these recent findings support a qualitative distinction between cortically based and subcortical levels of information processing. They imply that not all emotional responses are mediated cortically; rather, some may be initiated without any cognitive participation: “Emotional responses can occur without the involvement of the higher processing systems of the brain, systems believed to be involved in thinking, reasoning, and consciousness” (LeDoux, 1998, pp. 161)

Our Model of Emotion Regulation includes as well two levels of processing in relation to the aforementioned concept of processing but we consider that these two levels are connected closely with each other and that information is processed not only in a serial way but also concurrently. The experiential level includes the notions of emotional experience and emotional expression. Emotional experience is the covert emotional condition that a human is experiencing as a result of a stimulus or information of such kind. Emotional expression is the overt reaction of such a stimulus, the behavior that follows the experience. On the other hand, the rational level is comprised of the notions of emotion recognition, emotional management and emotional motivation. Emotion recognition is the ability to realize the true nature of an emotion as it is and to feel it in the appropriate degree. Emotional management is the ability to manipulate and to control an emotion while emotional motivation is the ability to transform an emotional experience into a motivational

urge. A visual representation of our model can be seen in figure 1. We believe that these two systems can interact. If someone during the stage of emotion recognition realizes intuitively that the emotion that is about to be triggered will have a negative and unpleasant emotional experience as an outcome, then it will be implicitly transformed to a different emotion so that it will be easily manageable in the next stage. The human brain prioritizes based on the principles of self-regulation and not on the search of objectivity and truthfulness.

### **3. THE CONCEPTS OF EMOTIONAL EXPERIENCE AND EMOTIONAL EXPRESSION**

The study of emotional experience and emotional expression has a long history, which dates back to the 1870s with scientific investigations undergone by Charles Darwin (Darwin, 1872). Darwin's work emphasized the biological utility of emotional expression. Thus, it contributed to the development of an evolutionary-expressive approach to emotion, which suggests that emotion exists because it contributes to survival (Oatley, 1992). Emotional experience, emotional expression and emotional arousal have been conceptualized as three primary components of emotion (Kennedy-Moore & Watson, 1999), with emotional reflection as a secondary component, involving thoughts about the three primary components.

Our model of emotion regulation distinguishes mechanisms surrounding the experience of emotions, from those surrounding the expression of emotions. Whilst in practical terms this is probably a seamless process, we believe it is conceptually useful to distinguish experience from expression. We hypothesize that it is more fundamental and harmful to control emotional experience, than to control emotional expression. The expression of emotions is behavioral. Thus the mechanisms surrounding it, involve the real and imagined consequences of expression, cultural and family rules for acceptable expression. These mechanisms may be different from those involved in emotional experience, which is of course experiential, rather than overtly behavioral. Such emotional experience may involve feeling too much intensive emotion, feeling inappropriate emotion, or feeling numb. Also important, is how the initial negative stimulus is registered, whether emotions are experienced as a gestalt, rather than separate somatic constituents and understanding the causes and meaning of the emotional experience. In short, it could be said that emotional experience points more towards a stimulus event, and expression more towards the behavioral response.

In summary, emotion regulation is not so much concerned about whether emotional expression is right or wrong but more with what mechanisms underlie successful and unsuccessful processing. Failure to express emotions may be integrally related to failure to properly process an emotional event. However, this is only one important part within a more complex process, as emotion regulation is regarded as the overall concept within which, emotional expression simply constitutes the final stage.

### **4. EXPERIMENTAL EVALUATION**

In this first experimental stage we wanted to investigate the implications behind the first level of emotion regulation and see how emotional experience and emotional expression interact with decision making and problem solving styles. Decision making and problem solving are two processes that circle almost every aspect of human activity. This way we can find some implications that connect emotion and its reactive responses with behavior in other areas that can be implemented in web design in various fields like e-learning, e-assessment and e-commerce. We hypothesized that highly emotional human beings will have a tendency towards emotional styles and not rational ones. Respectively this information can be used in web design to personalize content and navigation to their likings. For example a user that as a decision maker is dependent (does not like to decide on his own, values the advice of others) will enjoy a more solid, concrete and "closed" navigational system and not a web interface with many links and freedom of navigation or will opt for help and guidance more often than someone who is not dependent and likes to decide always on his own.

## 4.1 Sampling and Procedure

The study was carried out within one week and the participants were all Greek citizens that live in Greece. All participants were of relatively young age studying or working at the time of administration. They could either participate in the experimental sessions that were held in the New Technologies laboratory in University of Athens or fill in the questionnaires that could also be found online in the web page designed specifically for that purpose. They were all given a battery of questionnaires. A total of 247 questionnaires were completed and returned. 55 of them were half completed or had double answers and were omitted from the sample. Our final sample included 192 participants giving a completion rate of almost 80%.

Participants varied from the age of 18 to the age of 40, with a mean age of 27 and a standard deviation of 5. 73 respondents were male and 119 were female. Among other demographic characteristics that were examined were the profession and the computer experience level of each participant.

## 4.2 Questionnaires

The study used questionnaires to collect quantitative data. It included five measures, one each for personality, emotional arousal, emotion regulation, decision making styles and problem solving styles. Our first treatment involved the close examination of the experiential level of the emotion regulation questionnaire (emotional experience and emotional expression) and its correlation with decision making and problem solving styles. To evaluate Decision Making we used the General Decision-Making Style Inventory (DMSI) by Scott and Bruce (1995) which includes 25 items and 5 scales (Spontaneous, Dependent, Rational, Avoidant, Intuitive) and for Problem Solving the Problem Solving Styles Questionnaire (PSSQ) by Parker with 20 items and four scales (Sensing, Intuitive, Feeling, Thinking).

## 4.3 Design

Internal consistency was assessed by computing Cronbach alphas for the three measures. Although there are no standard guidelines available on appropriate magnitude for the coefficient, in practice, an alpha greater than 0.60 is considered reasonable in psychological research (Kline, 2000). After the inspection of the alpha coefficients, we performed descriptive statistics for the study sample as a whole and for the particular scales under investigation to examine the sample's suitability. Since our sample was normally distributed with variance of suitable proportions we continued our statistical analysis with the use of the statistical package SPSS. The statistical analysis used to perform this study was mainly one-way Analysis of Variance (ANOVA). Our research hypothesis stated that the experiential emotion regulation factors will have an effect on the participant's style of action. More specifically, participants that score high in emotional experience and emotional expression scales will have a tendency towards more emotional and less rational styles.

## 5. RESULTS

For the purposes of the experiment, Analyses of Variance (ANOVA) were performed in order to indicate the relationships between the variables of the study. Table 1 presents the main findings between the scale of emotional experience and the scales of the DMSI and PSSQ. The analyses indicated that emotional experience correlated highly with the spontaneous, rational and avoidant styles of the decision making questionnaire and the feeling and thinking styles of the problem solving questionnaire.

Table 1. Statistical Significance between the Emotional Experience scale and Decision-Making and Problem-Solving Styles

Constuct	F	Sig.
DM-Spontaneous	18.160	.000**
DM-Rational	7.907	.005*
DM-Avoidant	10.116	.002*
DM-Intuitive	14.469	.000**
PS-Feeling	33.562	.000**
PS-Thinking	11.025	.001**

\*p<0.005      \*\*p<0.001

A person that experiences emotions vividly is typically afraid that he might feel anxious, tense and moody. He can get emotional easily and therefore is reasonable to react in a spontaneous and not thoughtful way in occasions or with an inhibition of action in others. His pattern of behavior is tense as his character and is subjective to strong feelings. On the other hand a less emotional individual is more rational and more methodical in his behavior.

The exact same pattern is repeated with the emotional expression scale as it can be seen in table 2. This is consistent with the idea that since expression is the consequence of experience it will follow the same set of rules that govern experience. In the general population a person that experiences an emotion of a specific magnitude will have a reaction of equivalent proportions.

Table 2. Statistical Significance between the Emotional Expression scale and Decision-Making and Problem-Solving Styles

Constuct	F	Sig.
DM-Spontaneous	18.033	.000**
DM-Rational	18.090	.000**
DM-Avoidant	12.155	.001**
DM-Intuitive	7.077	.008*
PS-Feeling	19.469	.000**
PS-Thinking	19.189	.000**

\*p<0.005      \*\*p<0.001

In figure 2 we can see the means of both measures in all decision making and problem solving styles. The logical assumption is that the two notions of emotional experience and emotional expression will be highly correlated which indeed is the case. Pearson’s r has shown a significance at the 0.01 level (two-tailed) of .626.

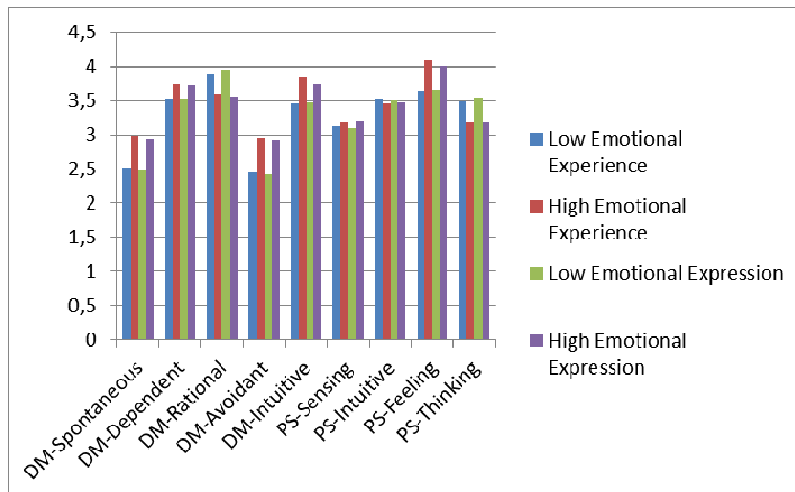


Figure 2. Means of High and Low Participant Groups in Emotional Experience and Emotional Expression. Emotional participants have higher means in the more “emotional” styles of spontaneous, avoidant, intuitive and feeling while less emotional participants score higher in the “logical” ones such as rational and thinking.

## 6. DISCUSSION

It may come as no surprise that emotional factors are important in the decision and problem solving process. The emotion regulation factors comprise characteristics that people often exhibit in their decision making. Apart from the standard emotion regulation questionnaire we developed a theory and a corresponding battery of questionnaires for the concept of Affect (Lekkas et al, 2009). The next step of our research is to combine these findings with the purely affective elements of our model. It has been argued that positive affect increases motivation, attention, pleasantness, participation and engagement, while negative affect is highly involved with boredom, fear, anger, displeasure and distraction.

By combining the personality style and the affective state of the individual with his regulatory mechanism (experiential and rational emotion regulation) we can reach into a conclusion of how affect influences his performance and the outcome of his behavior. At the same time our level of implementation after analysing our findings in decision making and problem solving preferences, will concentrate directly on the user learning process. We have already developed a web system based on learning performance evaluation for the testing of the various instruments that we have incorporated in our model (Germanakos et al, 2007). The cognitive elements are more straightforward since they are easier to measure and easier to quantify and we have already reached a level in which we can make inferences about how users with different cognitive abilities and preferences can be aided or guided through a personalized web interface (Tsianos et al, 2008). The final step to complete the implementation of our model is to add the affective elements and to investigate the inner and deeper relations that exist between them. Personality type is also a fundamental construct since personality research is already established and developed to a great extent. Our next task is to examine our findings in combination with the constructs of task-specific anxiety and personality.

Emotional and decision factors can be proven significant in defining user behavior in web applications and interfaces, taking into consideration psychometric challenges, as well as the complicated matter of quantifying and subsequently mapping emotions on a digital environment. Most theories of choice assume that decisions derive from an assessment of the future outcomes of various options and alternatives through some type of cost-benefit analyses. The influence of emotions on decision-making is largely ignored. The studies of decision-making in neurological patients who can no longer process emotional information normally suggest that people make judgments not only by evaluating the consequences and their probability of occurring, but also and even sometimes primarily at a gut or emotional level (Damasio, 1994).

Decision-making and problem solving are cognitive processes where the outcome is a choice between alternatives. We often have different preferences as to our approach, varying between thinking and feeling. When we use reason to make decisions, we seek to exclude emotions, using only rational methods, and perhaps even mathematical tools although emotions exist in the first stage of our decision making procedure and are followed by reasoning. The foundation of such decisions is the principle of utility, whereby the value of each option is assessed by assigning criteria (often weighted). Web systems until recently tried to integrate tools that aid user in a purely rational process (e-learning and decision-support systems). There is a whole range of decision-making that uses emotion, depending on the degree of reason that is included in the process. A totally emotional decision is typically very fast. This is because it takes time (at least 0.1 seconds) for the rational cortex to get going. This is the reactive (and largely subconscious) decision-making that you encounter in heated arguments or when faced with immediate danger. User Behavior is in its final analysis a decision making process. The nature of its activity is strongly correlated with emotions, that is why the role of emotions is extremely important in a setting like this. The mediating role of technology can help the designers to understand the emotional mechanisms of the users and adjust more efficiently to their needs.

One possible implementation of a Web-based system's interface that can appraise human emotion is through the use of a set of parameters that can adapt according to the emotional condition of the user and his preferred style of action. An emotionally tense or unstable individual will be able to adjust the contents of a webpage based to what he considers easier to control and manipulate. A certain emotional condition demands a personalization of equivalent proportions. The user will have the capability to respond emotionally either after being asked at a specific moment or after an initial profile construction.

Such a system should be designed in a way that it can create a detailed profile for every user and can provide two basic services. One application-based that will have to do with the interface, the navigation and its usability and aesthetical appearance and one content-based that will have to do with the database, the allocation of content, the depth and the dissemination of information. Using these, the interface will take the form that the user wishes so that he can work there more efficiently and less anxiously. Researching on decision making and problem solving is only the first step to map and model user patterns of behavior. The research results can be further used as more specific design guidelines.

## REFERENCES

- Bechara, A. et al, 2000. Emotion, decision-making, and the orbitofrontal cortex. *Cerebral Cortex*, vol. 10, pp295–307.
- Cingil I. et al, 2000. A broader approach to personalization, *Communications of the ACM*, Vol. 43, No. 8.
- Damasio, A. R. 1994. *Descartes' error: Emotion, reason, and the human brain*. Putnam Publishing Group, New York.
- Darwin, C., 1872. *The Expression of the Emotions in Man and Animals*. D. Appleton and Company, New York.
- Eklund J. and Sinclair K., 2000. An empirical appraisal of the effectiveness of adaptive interfaces of instructional systems. *Educational Technology and Society*, Vol. 3, No. 4.
- Germanakos, P. et al, 2007. Capturing Essential Intrinsic User Behaviour Values for the Design of Comprehensive Web-based Personalized Environments, *Computers in Human Behavior*, Special Issue on Integration of Human Factors in Networked Computing.
- Kennedy-Moore, E. and Watson, J. C., 1999. *Expressing Emotion. Myths, Realities, and Therapeutic Strategies*. The Guildford Press.
- Kim W., 2002. Personalization: Definition, Status, and Challenges Ahead, Zurich, *Chair of Software Engineering JOT*, Vol. 1, No. 1.
- Kline, P., 2000. *Handbook of Psychological Testing*. Routledge, London.
- LeDoux, J., 1998. *The emotional brain*. Touchstone, New York.
- Lekkas Z. et al, 2009. Implementing Affect Parameters in Personalized Web-based Design, *Proceedings of the 13th International Conference on Human-Computer Interaction – HCI International 2009 (HCI 2009)*, pp. 320-329, LNCS & LNAI, Springer-Verlag Berlin Heidelberg San Diego, CA, USA.
- Lewis, M. and Haviland-Jones J.M., 2004. *Handbook of emotions (2nd Ed.)*. The Guildford Press, New York.
- Mulvenna, M. D. et al, 2002. Personalization on the net using Web mining, *Communications of the ACM*, Vol. 43, No. 8, pp 123–125.
- Norman, D., 2004. *Emotional Design: Why We Love (or Hate) Everyday Things*. Basic Books.
- Oatley, K., 1992. Integrative action of narrative. In D. J. Stein & J. E. Young (Eds.) *Cognitive science and clinical disorders*. Academic Press, San Diego.
- Picard, R.W., 1997. *Affective Computing*. MIT Press, Cambridge.
- Samoilov, A. and Goldfried, M.R., 2000. Role of emotion in cognitive-behaviour therapy. *Clinical Psychology, Science and Practice*, Vol. 7, No. 4, pp 373-383.
- Scott, S. G. and Bruce, R. A., 1995. Decision-making style: the development and assessment of a new measure. *Educational and Psychological Measurement*, Vol. 55, No. 5, pp 818-831.
- Tsianos, N. et al, 2008. User-centered Profiling on the basis of Cognitive and Emotional Characteristics: An Empirical Study, *Proceedings of the 5th International Conference on Adaptive Hypermedia and Adaptive Web-based Systems (AH 2008)*, pp. 214-223. Springer-Verlag Berlin Heidelberg, LNCS 5149.

# A STUDY OF A NAVIGATION SYSTEM THAT INDUCES TOURISTS TO VISIT SIGHTSEEING SPOTS AGAIN VIA A FEELING OF REGRET

Masaki Masuda<sup>1</sup> and Yoshio Nakatani<sup>2</sup>

<sup>1</sup>*Graduate School of Science and Engineering*

<sup>2</sup>*College of Information Science and Engineering*

*Ritsumeikan University, 1-1-1 Noji-Higashi, Kusatsu, Shiga, JAPAN*

## ABSTRACT

Increasing the tourists is important at sightseeing spots. There are two types of reasons people visit sightseeing spots again: Because they are particularly interesting or something interesting was overlooked. The latter is based on a “feeling of regret.” The feeling of regret can actually change people’s behavior. The feeling of regret involves the difference between the ideal and reality, and the difference between the effort they put in and whether if feels successful or not. In psychology the Zeigarnik effect [1] states that people remember “unfinished or interrupted” tasks better than those they have completed. With sightseeing we consider the Zeigarnik effect to be involved in something a person wished to view or visit but couldn’t. We predict the effect could be used induce people to want another chance, or that is, we think that they will wish to visit the relevant sightseeing spot again. The system proposed in this paper provides pictures and information in a gradual manner. In this paper we suggest how to get people to visit sightseeing spots again through the creation of “their sightseeing having been incomplete”. An experiment to evaluate the proposed prototype system was performed in Kyoto. The evaluation experiment took place in the middle of January in 2011. As a result, the test subjects were felt a regret by providing pictures using a temporal difference and a spatial difference, we confirm the efficacy of the study.

## KEYWORDS

Zeigarnik effect, Sightseeing, Navigation system

## 1. INTRODUCTION

### 1.1 Background

Sightseeing is a global business and forms a big market in Japan. According to the World Tourism Organization (WTO) the number of global tourists totaled 25 million people in 1950, 69 million in 1960, 159 million in 1970, 287 million in 1980, 455 million in 1990, and 700 million in 2000. It is now being predicted that the tourist market will increase from here on in [2]. Sightseeing has been gaining attention in Japan because it will increase consumption and employment. The “Tourism-based Country Promotion Basic Act” was enforced on January 2007, with the Japan Tourism Agency having been organized to materialize a Tourism-based Country. This is being evaluated as being the pillar of a very important policy, or that is, a concern about sightseeing has been growing in recent years both globally and in Japan.

Package tours made available by travel agencies are widely used but with tourists deciding their own routes and plans and not using package tours having grown in popularity in recent years. That type of sightseeing is considered to be “Autonomous Sightseeing.” In recent years people deciding a travel plan using information they have accessed has led to the expectation that “Autonomous Sightseeing” rather than “Heteronomous Sightseeing” will increase in popularity [3]. Development of information service on the internet promotes this tendency. Short tours are being focused upon by young businessmen and women, with an increase in the number of people that enjoy the planning process.

The sightseeing scene in recent year has thus been changing, therefore making a direct connection between tourists and sightseeing spots important.



The common point of sightseeing resources has the feature that “timing” of a visit raises more charm of the sightseeing element. Natural landscape changing every season and the event and the festival which are held only at decided time are the sightseeing resources with the high rarity more than the historical building and the streets which can be seen anytime. “Getting rare timing” is getting the interest of a repeater tourist. So, the timely event induces revisit as a repeater in a sightseeing spot and timely provision of events information is important. It is desired to use a tool using a sightseeing navigator and devise how to show it a sightseeing element as well as a device on the host side of the tourist spot [4].

## 1.2 Related Research

### 1.2.1 Sightseeing Navigation System Based on Benefit of Inconvenience

Systems used to cut down on labor have been actively developed with the rise of technology in recent years. However, those technologies have also given rise to several problems as greater convenience has a negative side to it too. For example, face to face encounters are lost when people meet each other over the internet. The “benefit of inconvenience” is thus attracting attention. If we regard convenience to be less work we can also discover other benefits beside doing away with grudge work [5].

Several studies were made on sightseeing navigation systems that are based on the benefit of inconvenience by Ken Tanaka and Kanako Ichikawa. These studies emphasize a “sense of fun” and suggest a tourist navigation system that intentionally limits route information.

Confirmation took place on what interactions with the environment should be promoted by Kanako Ichikawa drawing up a sightseeing plan using an electronic map before the actual sightseeing took place, and deleting the electronic map on the sightseeing day, then only showing a rough sightseeing plan and their current position via GPS (Global Positioning System) (Figure 1) [6]. This induces accidental encounters occurring.

Ken Tanaka’s study does not support the most efficient sightseeing and instead sightseeing that can be leisurely enjoyed. In this system, the map around an user disappears as the user moves around. The effectiveness of it was then tested in assessment experiments utilizing a subject (Figure 2) [7].

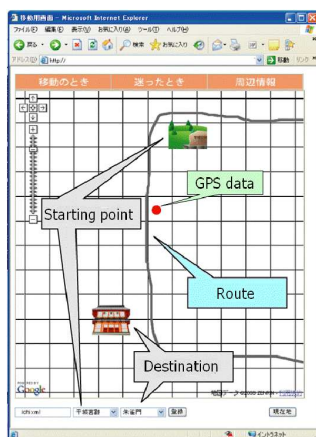


Figure 1. Tourist Navigation System (Ichikawa)

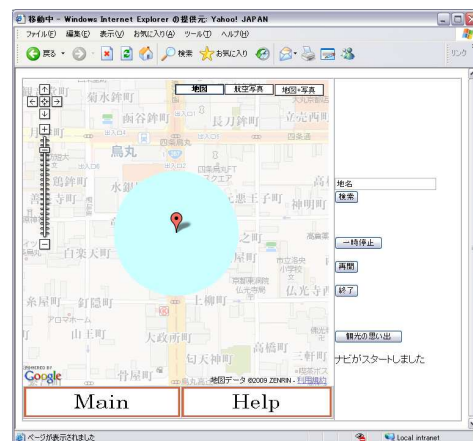


Figure 2. Tourist Navigation System (Tanaka)

### 1.2.2 Zeigarnik Effect

People have a variety of feelings, with a particular feeling that tends to be related to behavior and remain for a long period of time being a “feeling of regret.” The feeling of regret can actually change people’s behavior. The feeling of regret involves the difference between the ideal and reality, and the difference between the effort they put in and whether if feels successful or not. It is a given fact that being serious about a task is related to the next action people take in such a situation. We often experience a feeling of regret because of the seriousness we consider a task to have [1].

The influence that regret has on people in psychology and behavior was studied by the Russian psychologist Bluma Zeigarnik. She discovered that unfinished or interrupted actions remain in the memory better than those completed.

With sightseeing we consider the Zeigarnik effect to be involved in something a person wished to view or visit but couldn't. We predict the effect could be used induce people to want another chance, or that is, we think that they will wish to visit the relevant sightseeing spot again.

Some studies have focused on the benefit of inconvenience but our study considers not only the benefit of inconvenience but also the Zeigarnik effect.

## 2. SYSTEM ARCHITECTURE

### 2.1 Approach of this Study

A number of efficient sightseeing navigation systems have been developed for use in the field of sightseeing but a system to induce people to visit sightseeing spots again has yet to have been developed, which needs to include the following:

- 1) Support for chances and discoveries based on the benefit of inconvenience
- 2) Induction of tourists' revisit to sightseeing spots

We have developed a system incorporating the above.

Anyone walking around sightseeing spots can possibly make new discoveries. The system in this study allows tourists to imagine the enjoyable things they could discover if they were to visit a spot at the right time by removing times and seasons from the picture. When tourists arrive at a goal the system displays pictures that they can't see at the sightseeing spots on the map, thus inducing a feeling of regret in the users. They would then think "I want to come here again." They cannot fulfill the imagined sightseeing because of the different time and space, or that is, it remains "unfinished sightseeing." They feel regret about the unfinished experience, and will thus probably wish to visit to the sightseeing spot again.

### 2.2 System

This study involves obtaining GPS data and providing information this is only viewable on sight via utilization of iPhones. The application language used was Objective-C. We considered the language to suit our system because it includes a method of setting the distance between the present location and the goal.

An outline of the system is provided below.

1) GPS data is updated at equal distances. When the user approaches the destination, which includes registered picture information, the system provides a picture of the object of interest but shifts in time slightly in some pictures.

2) The users attempt to find the place the picture depicts, and can often be near it. They search for the destination by checking the distance between their present location and the destination via use of changing colors. The color changes from "blue-yellow-green-red" over distance.

3) The same scenery as the picture is impossible to view because of a few differences between the most suitable time and the present exist with the same spot as in the picture. The system thus induces a feeling of regret.

4) After the user arrives at the destination pictures of sightseeing spots they did not visit but were near are displayed on a map based on GPS data. If they have no time to return to that place the system again induces a feeling of regret.

#### 2.2.1 Main Screen

The main screen displays pictures that have been shifted in time slightly. If the system shows pictures they are interested in they cannot see the same view as in the picture, however, thus leading to a feeling of regret. The important point is that the time is slightly different. If the picture's timeframe has been shifted too much users tend to feel resignation rather than regret. The important point is what degree of time slip produces the

most regret. We propose the hypothesis below because we have no psychographic or cognitive science knowledge.

- 1) In the case of a seasonal event the previous or following season's picture is displayed. For example, if spring is the best season early spring (end of winter) or early summer.
- 2) In the case of a weekly event the previous or following week's pictures is displayed.
- 3) In the case of a daily event the previous or following day's picture is displayed.
- 4) We imagine that the picture of the previous time's event is better than the following in inducing a feeling of regret.

The system displays pictures selected based on the above standards.

The main screen displays GPS data, buttons, pictures, and colors (Figure 3). Users can select an object of interest using the buttons. The system changes behavior when objects of interest are selected. The reason that users need to select objects of interest is that places they are not interested will not induce a feeling of regret. The system then utilizes GPS data and the objects of interest to display pictures when the user is near the destination. Users look for the same place, survey the scenery, and then enjoy viewing the scenery. When the users arrive at the destination they notice a difference between the picture and the sightseeing spot they are viewing, leading to a feeling of regret.

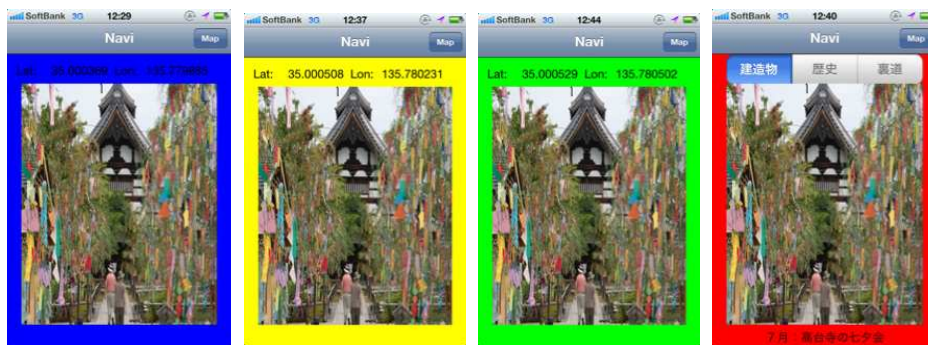


Figure 3. Main screen

Over the past few years several studies have been made on sightseeing navigation systems that are based on the benefit of inconvenience. The studies then revealed the problem of a feeling of insecurity.

This study therefore depicts the distance from the present position to the goal using colors (Figure 3). When the distance is 100 meters the system shows a picture within a blue frame, as in Figure 3. The colors then change through the sequence of “blue-yellow-green-red” as users approach the destination. The distance from the present position to the destination uses colors because colors are more intuitive than numbers, and we wish to enable users to feel a sense of pleasure.

### 2.2.2 Sub Screen

The sub screen displays a map and information on a sightseeing spot they did not visit but were near to (Figure 4).

The sub screen also shows their route using red pins on an electronic map.

The system utilizes GPS data and any objects of interest to display trackside sightseeing spots that the users did not visit using green pins on the electronic map. The system displays a popup window if the users touch the green pin through which they gain information on that sightseeing spot. The system shows a picture of the sightseeing spot if users then touch the detail-button in the popup window. The “Trackside sightseeing spots” in this study were set to be within 100 meters of the user's route. 100 meters was considered the most suitable in enabling the users to feel “I could have deviated from my route and dropped in at the spot, but didn't.” We predict that a feeling of regret will be induced in users from this feature.

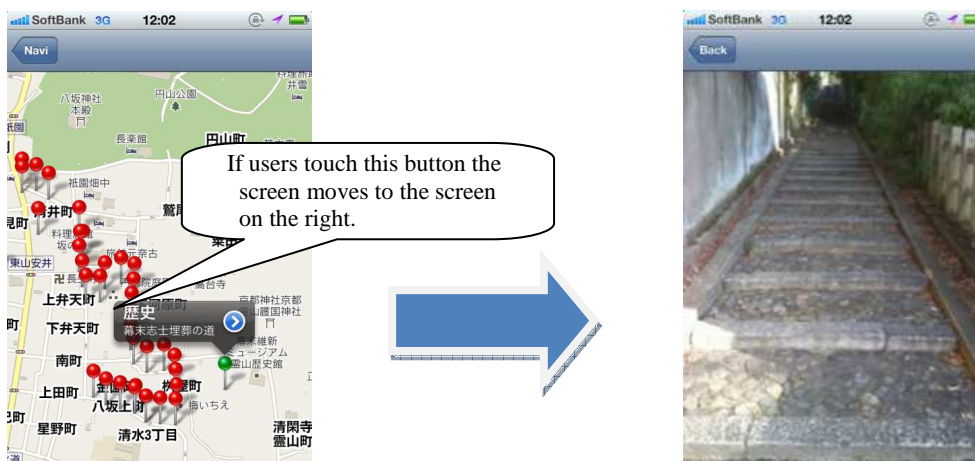


Figure 4. Sub screen

The green pin only displays once the users have arrived at the destination shown on the main screen. If the users do not reach the destination they will not be able to view the trackside sightseeing spots they did not visit because the green pin will vanish.

The system enables the situation where users can't look for the spot by removing the information on it, thus also inducing a feeling of regret.

### 3. EVALUATION

An experiment to evaluate the proposed system was performed on 7 subjects. The evaluation experiment took place in the middle of January in 2011, and is described below.

#### 3.1 Experiment

An area around Kiyomizu temple in Higashiyama-ku, Kyoto, was used in the experiment.

The test subjects were 7 people of their 20's who live in the Kansai area. They were divided into three groups on the basis of the objects of interest (Table 1). They selected two subjects from history, buildings, and back roads. Table 1 provides information on objects of interest.

Table 1. Grouping and the attribute of the test subjects

	Group 1	Group 2	Group 3
Number of people	2	3	2
Object of interest	Buildings	History	Back roads
Ratio of males to females (men: women)	1:1	1:2	1:1

The test subjects went sightseeing equipped with iPhones. We went with them to thus be able to observe them. When they arrived at the goal we directed them to view the sub screen. The sub screen displayed information on a sightseeing spot they did not visit but were near.

After the sightseeing the test subjects were provided with questionnaires.

#### 3.2 Experimental Results

The 3 groups enjoyed going sightseeing. Figures 5, 6, and 7 show the routes of each group. The yellow circle in the picture marks where the picture has been slightly shifted in time. The purple circle in the picture marks the goal. Figures 8, 9, and 10 show pictures with slight time shifts. Figure 8 shows a picture half a year later, Figure 9 1 month later, and Figure 10 1 month prior.



Figure 5. Route of Group 1



Figure 8. Picture slightly shifted in time (buildings)

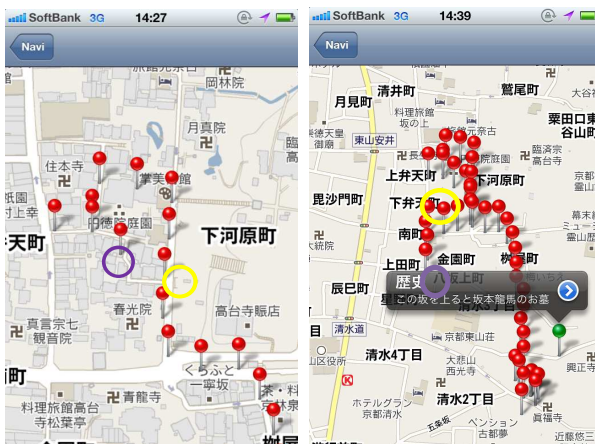


Figure 6. Route of Group 2



Figure 9. Picture slightly shifted in time (history)



Figure 7. Route of Group 3



Figure 10. Picture slightly shifted in time (back roads)

### 3.2.1 Considerations with Benefit of Inconvenience

The difference between the current position and destination was depicted using colors the system navigation based on the benefit of inconvenience. We predicted that most people would prefer to view the environment this way.

We took a look lost on way on Figures 5, 6, and 7. Interaction can be expected which asked a person the current way and look around attentively to be lost [7].

We asked whether the time to view the system screen while moving around or sightseeing using a map and tour guide would take more time the bottom, which was then judged using an ascending scale of 1 to 5 (short-1, long-5), leading to the discovery that using the system took less time than using a map and a tour guide. One of the reasons for that was it proved easier to understand the connection between the current position and the goal when depicted with colors, which led to less time being required.

Moreover, they also answered a questionnaire on how it felt to search for a destination using only a picture on an ascending scale of 1 to 5 and a “feeling of anxiety”, “level of difficulty”, and “amusingness”, with the averages being provided below.

Feeling of anxiety (anxiety-5, reassurance-1)	3.28
Level of difficulty (difficult-5, easy-1)	3.85
Amusingness (amusing-5, not amusing-1)	4.85

The value of "Amusingness" was high at 4.85, although feelings of anxiety and difficulty were also present. This then led to the consideration of why it was possible to feel "amusement" but also anxiety and difficult. To the question “How did it feel searching for a destination using a picture and the distance?” they answered “I didn't know the way, but it was fun to search for it,” “I slightly detoured but still got there,” and “It's fun because it is introduced an element of chance.” They thus presumably enjoyed the element of chance which incorporated anxiety and the opportunity to make an accidental discovery. The answer about making an impression on them in this sightseeing were as in the sunset in Kyoto which was found near the goal and the building and the lane they did not know so far. These are the encounters with “accidental discoveries” which can't be met for conventional sightseeing. It can be said a destination using only a picture and colors induced an “accidental discovery” and “amusement” which can go sightseeing as thought it were a game.

### 3.2.2 Consideration about the Degree of "regret" from the Temporal Difference

The time of the pictures of 3 the destinations were shifted:

- To one month later (history)
- To one month prior (back roads)
- To half a year later (buildings)

\*Inside the () gives the objects of interest.

In this way we examined the resulting feeling from the test objects being shifted in time. We questioned them on the degree of regret the objects of interest they gone to induced, with the answers being provided below.

Table 2. Destination and number of people that answered made an impression

	Objects of interest	Number of people that answered made an impression
Group 1	History	1
	Buildings	1
Group 2	Back roads	3
	History	0
Group 3	Buildings	0
	Back roads	2

Before the experiment we assumed that a later event would induce a feeling of regret, but we could not define a causal relationship. The reason for that is because the impact of a picture was stronger than the impact of timeframe being slightly shifted.

However, the answers made us more aware of how a shift in time has an effect. The answers included “If I came here a month ago I could have seen the Hana-touro,” and “I should have come here a little earlier.” All the test subjects said they wished to visit there again, thus affirming that the picture being shifted in time resulted in a feeling of regret.

### 3.2.3 Consideration about the Validity of A Spatial Difference

The system displays a map and information of a sightseeing spot they did not visit and but were near to only after they have arrived at the destination. We requested their impressions of the spot near to the test objects.

They explicitly answered that the spot left them with an impression, which is considered to trigger another sightseeing excursion. Many of them stated that they would to have stopped at the spot if they had more time.

The most important problem to induce them to think about visiting again by ensuring they did not complete the sightseeing. Group 1 couldn't find the spot, even though they did search for it. They very clearly mentioned they felt regretful that they couldn't find it, thus leading to the prediction that regret of not having found a spot would trigger sightseeing there again. The timing and display time thus suit inducing a feeling of regret, and we can be sure that consideration of the validity of a spatial difference exists.

#### **4. FUTURE WORKS**

Various problems were discovered in this study, including the difficulty of ascertaining the feeling induced from the shift in time with the test objects, which will be studied from here on in. We will also add a vibration function to the system so that the change in color can be understood to have occurred even if the screen is not being viewed. We think the system can be a useful tool for use in connecting sightseeing spots and tourists over a network.

#### **ACKNOWLEDGEMENTS**

Special thanks go to the research students of Nakatani laboratory who provided such invaluable comments and warm encouragement.

#### **REFERENCES**

- [1] Zeigarniku, B. , Über das Behalten erledigter und unerledigter Handlungen, *Psychologische Forschung*, 9,1927.
- [2] The Tourism Agency, 2011. About The Tourism Agency. <http://www.mlit.go.jp/kankocho/about/setsuritsu.html>
- [3] Syuzo Ishimori, 2001. The Potentialities of Autonomous Tourism in the Twenty-first Century, *Senri Ethnological Reports*, pp.5-14 (in Japanese)
- [4] Kaoru Okamura, 2007. Mototsugu Toyonaka, 2007. How to Promote Repeaters? Empirical Analysis of Tourist Survey Data in Kansai Region, *KISER Discussion Paper Series*, No.10. (in Japanese)
- [5] Hiroshi Kawakami, et al., 2009. System Design based on Benefit of Inconvenience and Emotion. *ICROS-SICE International Joint Conference 2009*, pp.1184-1188.
- [6] Yoshio Nakatani, Ken Tanaka, and Kanako Ichikawa, 2010. A Tourist Navigation System that Promotes Interaction with Environment. *IAENG International Journal of Computer Science*, Vol.18, No.2, pp.1-8.
- [7] Ken Tanaka, 2010. Yoshio Nakatani, 2010. Sightseeing Navigation System that Promotes Interactions with Environment by Restricting Information, *IEEE SMC 2010*, pp.453-458.

# PROCESS-INDUCED DECISION COSTS ON SEQUENTIAL VALUE JUDGMENTS

Mei-Chun Wu<sup>1</sup> and Feng-Yang Kuo<sup>2</sup>

<sup>1</sup>*Department of Information Management, Nanhua University  
No. 55, Sec 1, Nanhua Road, Chiayi, Taiwan*

<sup>2</sup>*Department of Information Management, National Sun Yat-sen University  
70 Lien-Hai Road, Kaohsiung, Taiwan*

## ABSTRACT

In repeated-choice situations, people tend to stick to the previously chosen alternative in their subsequent decision. While “effort-as-information” and “resource availability” produce similar results, the manner of resource expenditure involves different coping strategies in subsequent decisions. We investigated the impact of process-induced decision costs of previous decision on subsequent decision. Lower consistency rate occurred when additional resources caused by layout change were required. The decreased consistency rate implies that resource availability play a significant role in sequential decision-making situations. Further, making a difficult preliminary decision (as reflected by longer response times) can deplete self-regulation resources, producing a higher likelihood of a decision inconsistency when fluent processing was impeded by layout change. The research findings suggest that the popular use of dynamic web pages in online shopping situations is likely to increase processing costs by changing product locations which may potentially influence consumer judgments. Both consumers and managers should be aware of such underestimated effects.

## KEYWORDS

Decision costs, value judgment, effort-as-information, resource availability.

## 1. INTRODUCTION

The choices are often made repeatedly, rather than made isolated from previous choices. In repeated-choice situations, consumers' prior choices have been shown to impact their current choice processes and outcomes (Chen and Rao 2002; Monga and Rao 2006; Thaler and Johnson 1990). According to the explanations of “effort-as-information” (Arkes and Blumer 1985; Kruger et al 2004; Loewenstein and Issacharoff 1994) and “resource availability” (Bettman et al 1998; Meyers-Levy and Tybout 1997), people will tend to stick to the previously chosen alternative in their subsequent decision. The concept of “effort-as-information” suggests that effort spending on initial decisions is deemed as a source of information for subsequent decisions. The other notion concerns “resource availability”. After depleting resources in initial decisions, people will use simple decision heuristics in making subsequent decisions. Although these two explanations produce similar results, the manner of resource expenditure involves different coping strategies in subsequent decisions.

Expending resources on a previous task has the potential to interfere with cognitive activities which could result in biased judgments (Vohs and Schmeichel 2003). However, existing research on consumer behavior usually attempts to find significant independent or moderating variables toward that choice. How the prior decision processes or outcomes influence subsequent decision-making has yet to produce a great deal of empirical research (Kim 2008). To gain further understanding of consumer decision-making, in this study we investigate the impact of process-induced decision costs of previous decision on subsequent decision.

According to literature review relating to repeated-choice, two weaknesses exist in the current research status. First, current research does not focus on the specific impact of the previous choice on subsequent ones. This research stream has failed to scrutinize the underlying mechanism of the impact of previous choice. Second, the existing research has failed to break down previous choice into a subordinate concept (e.g., process and outcomes). Existing studies have focused mainly on the choice outcomes themselves.



These studies ignore that the process (e.g., the amount of effort) of previous choices can also affect subsequent choices.

The objective of the present study was to investigate the accountability (i.e., “effort-as-information” or “resource availability”) of the impact of previous decision. We examined under what kind of situations and to what extent increasing processing costs (i.e., require more resources) alters the tendency of the subsequent decision to go with the previous decision. To control possible contaminations, the increased decision costs were generated by engaging in the processing activities themselves, rather than the costs associated with the information evaluation.

Specifically, the process-induced decision costs were manipulated by varying the required resources through changing the locations of objects that were seen in the first stage of judgment. As the objects were exchanged across two judgment stages, we expected the resources required to make the overall judgment to increase. Changing object locations increased the magnitude of processing effort that we were able to examine. Furthermore, when more resources were expended on the preceding judgment, changing object locations allowed us to test whether the proposed effects of processing difficulty on the subsequent judgment would be magnified.

Overall, this study investigated the influence of process-induced decision costs on sequential judgments. We expected that layout re-arrangement makes the judgment more demanding by increasing the cognitive workload and will influence the likelihood of the previously chosen alternative being selected. Additionally, we examined the interplay of information re-arrangement and decision costs expended in the preceding judgment in subsequent decision making behavior. Throughout, we attempted to address the accountability (“effort-as-information” or “resource availability”) of how the prior decision processes or outcomes influence subsequent decision making.

## **2. CONCEPTUAL BACKGROUND**

### **2.1 Processing Difficulty on Decision**

Judgments are influenced by experiences related to the mental effort (Schwarz and Clore 2006; von Helversen et al 2008). The notion that the process of processing may generate affect, in addition to affective reactions generated by processing the (conflict) information itself (Luce 1998), has gained an increasing attention in consumer behavior research (Garbarino and Edell 1997; Im et al 2010; Loewenstein 1996). The process-induced affect argues that negative affect can be elicited by a process that requires more deliberate thinking.

Process-induced negative affect by expending more cognitive effort was shown to influence choice of equivalent alternatives. Garbarino and Edell (1997) demonstrated that when people exerted more cognitive effort in processing an alternative, they experienced more negative affect. If the evaluations of the alternatives were equivalent, then the alternative that had less negative affect associated with it was chosen. The effort adversely affected choice of the more difficult to process alternatives, lowering the likelihood of difficult alternative being selected.

Processing difficulty due to visual presentation variables that impede fluent processing can produce deferral choices. In Novemsky et al (2007) study, consumers were presented with descriptions of two cordless phones and asked to choose the one they prefer, allowing them to defer choice if they had no clear preference. They found more than twice as many participants deferred choice when the font was difficult rather than easy to read. Also, Song and Schwarz (2008) demonstrated that the readability of a print font can have a profound impact on consumer judgment and choice. In their study, participants were provided with a description of an exercise routine, printed in an easy or difficult to read font. When the font was easy to read, participants reported higher willingness to make the exercise part of their daily life. In a second study, when the recipe was printed in a difficult to read font, participants inferred that preparing a Japanese lunch roll would require more effort and skill and were less inclined to prepare that dish at home. Throughout, the difficulty of information processing was mistaken as indicative of the difficulty of performing the described behaviors. These studies shed a light on that minor aspects of the visual display can significantly influence judgment and defer choice.

## 2.2 Trade-off vs. Dominance Decision

Making trade-off decisions requires more effort than that of making dominance decisions. That is, a decision-making involving a trade-off relationship requires more decision-related efforts or costs than one involving a dominance relationship. Quick response times point to dominance decisions where at least one of the alternatives is outstanding and slow response times point to trade-off decisions where the alternatives are equally attractive. For example, Klein and Yadav (1989) found that participants spend less time on decision-making when dominated alternatives were included. Luce (1998) found that in a high trade-off difficulty condition, decision-makers may choose to defer decision and avoid trade-off conflicts. Thus, dominance relationships provide decision-makers with an easy way of choosing among alternatives.

As environments require more cognitive effort to process information, decision makers often switch to decision heuristics. However, these heuristics may generate less accurate decisions, biased responses and preference reversals (Johnson et al 1988). Garbarino and Edell (1997) noted that people are willing to let go some benefits to conserve cognitive effort.

## 2.3 Effort-as-Information

The “effort-as-information” perspective suggests that after expending efforts, people attempt to preserve the decision outcome associated with previous effort in their subsequent tasks. Once an investment in money, time or effort has been made, people has greater tendency to continue an endeavor, termed escalation of commitment (Arkes and Blumer 1985). Several explanations for escalation of commitment include the desire not to appear wasteful (Arkes and Blumer 1985), the need to justify one’s previous decision (Brockner 1992; Staw 1981), and previous belief structure and involvement in the previous decision (Biyalogorsky et al 2006).

Expending resources in a previous decision promotes higher motivation to maintain resources by sticking with the preceding decision. Furthermore, decision difficulty increases the magnitude of maintaining one’s previous decision (Luce 1998; Samuelson and Zeckhauser 1988). Briefly, the previous decision process or commitment can influence the current decision by continuing or repeating the course of action. In a repeated-choice situation, people are more likely to retain their previous decision, due to the fact that a trade-off choice requires more effort than a dominance choice. As noted by Samuelson and Zeckhauser (1988, p. 37), “the larger the past resource investment in a decision, the greater the inclination to continue the commitment in subsequent decisions.”

Regarding the consequence of effort involving in the decision process, people have a tendency to use “effort” as a cue for their evaluations or judgments (Godek et al 2001; Kruger et al 2004). In Kruger et al. (2004) study, participants evaluated a poem more favorably when they thought that the poet took more time (i.e., 18 hours) to write the poem than when they thought the poet took less time (i.e., 4 hours). Godek et al (2001) showed that participants were happier with their choices and were willing to pay more for their chosen options when they made a choice with more effort than when they made a choice with less effort.

## 2.4 Resource Availability (The Cost of Decision-Making)

There are three different types of decision-related costs. Cognitive cost has been regarded as a basic cost of decision-making by many researchers (Bettman et al 1990; Shugan 1980). Emotional cost results from facing emotion-laden choices (Luce 1998). Trade-off difficulty can produce negative emotions. High trade-off difficulty (i.e., multiple goals cannot be achieved at the same time) produces highly negative emotions (Luce 1998).

Recently, researchers have proposed that choices are related to expending self-regulation resources. Self-regulation is defined as “the self exerting control to change its own responses in an attempt to pursue goals and standards” (Vohs and Baumeister 2004, p. 2). Self-regulation resources are limited (Baumeister and Heatherton 1996). Hence, performing one act of regulating the self can impair performance on a subsequent, apparently unrelated act of self-control.

Making a choice can deplete self-regulation resources, which then impairs the self’s ability to manage cognitive activity (Schmeichel et al 2003). In other words, the process of choosing can expend some resources, thereby leaving the executive functioning less capable of carrying out other activities. In Vohs et al

(2008) study, in the self-regulation-resource-depleted condition participants were instructed to make a binary choice between varieties of consumer products, such as magazines, colored pens, and t-shirts; in the self-regulation-resource-no-depleted condition participants were instructed to rate products. After the task, the participants were asked to drink as much of an ill-tasting beverage as they could. The results showed that participants making binary choices between several products drank fewer ounces of the ill-tasting beverage than those who merely rated the products. Vohs et al (2008) indicate that there is a hidden cost to choosing, which is different from merely thinking about options.

Although prior research (Schmeichel et al 2003; Vohs et al 2008) has shown that decision-making requires self-regulation resources, in those studies the subsequent tasks (e.g., drinking an ill-tasting beverage or practicing math problems) were to show the effect of the expenditure of self-regulation resources and not directly related to decision-making. Another important aspect of decision-related costs is that depleted resources cannot be restored immediately. Therefore, to study sequential decision-making situations, this aspect of decision-related costs must be taken into consideration.

### 3. RESEARCH HYPOTHESES

In this study, an alternative's overall value was a combination of the evaluation of its component objects. Respondents had to evaluate between two alternatives and choose the one with higher value in a two-stage value judgment task. We attempted to investigate, on exposure to two-alternatives-choice task, how consumer value judgments were influenced by process-induced decision costs that were generated in a more controlled manner.

To provide evidence for the explanation of "effort-as-information" versus "resource availability", we directly manipulated additional resource consumption in the middle of the first and second stage of value judgments. Specifically, after the first stage of value judgment, the component objects were rearranged either within the same alternative (within-swap) or between alternatives (between-swap). If effort expenditure or resource availability had a strong influence, it may play a role in consistent choice rates of sequential value judgments. The study focused on the additional efforts in the processing activities themselves, rather than the efforts associated with evaluating information, and the effect of this process-induced effort expenditure on value judgments.

This research attempted to investigate whether value judgments were altered by incremental processing difficulty. The logic behind this study was that if resource availability was at work, we should find a significant impact of additional resource expenditure manipulation on subsequent decision-making. Specifically, in the within-swap condition (i.e., component objects were rearranged within the same alternative after the first stage of value judgment), both the resource availability and effort-as-information explanations predict the consistent choice rate to be the same with that of no additional resource expending between the initial judgment and the subsequent one. However, in the between-swap condition (i.e., component objects were rearranged between alternatives after the first stage of value judgment), re-mapping of objects to alternatives generated processing costs. This additional resource expenditure was expected to influence the consistent choice rate. The resource availability explanation predicts the consistent choice rate of between-swap condition should be lower than that of no additional resource expending condition (no-swap). On the contrary, the effort-as-information explanation predicts the consistent choice rate of between-swap condition should be the same with that of no additional resource expending condition. Thus, we propose:

Hypothesis 1: When component objects were rearranged within alternatives, the consistent choice rate will be the same with that of no additional resource expending condition.

Hypothesis 2a: When component objects were rearranged between alternatives, the resource availability explanation predicts the consistent choice rate will be lower than that of no additional resource expending condition.

Hypothesis 2b: When component objects were rearranged between alternatives, the effort-as-information explanation predicts the consistent choice rate will be the same with that of no additional resource expending condition.

In trade-off situations where alternatives are equally comparable based on the evaluation of their component objects, decision-makers may devote more extensive efforts in evaluating objects, resulting in

longer response times. According to resource availability, such effort expenses in the initial value judgment may incur resource constraints and impair the self's ability to manage subsequent cognitive activity. Longer response times (i.e., more effortful processing) in the first judgment may interfere with subsequent judgment in the between-swap and within-swap conditions where additional resource expenses were required. Contrarily, based on the effort-as-information explanation there is no such impact of additional resource expenses on subsequent judgment. That is, there should be no difference in response times as a function of swap conditions. We propose:

Hypothesis 3a: The resource availability explanation predicts there is swap condition by consistent value judgment interaction on response time.

Hypothesis 3b: The effort-as-information explanation predicts there is no swap condition by consistent value judgment interaction on response time.

## 4. METHOD

### 4.1 Participants

Twenty undergraduate students at the University of Toronto Mississauga participated in the experiment. The participants were paid \$10 (Canadian) per hour.

### 4.2 Materials and Design

Stimuli were constructed using an image database containing 192 exemplars from each of 4 everyday object categories (hats, rings, bags and watches) for a total of 768 images. Several online shopping websites were used to extract these images. Each image displayed a product on white background and all images subtended 360 x 360 pixels. For each of the 4 product categories, 96 price-matched object pairs were created. As shown in Fig. 1, four object pairs, one from each category, were then combined to create the display sequence in each of the 96 experimental trials. Specifically, in each display, rows of four cells (each cell subtending 400 x 400 pixels) appeared on the top and the bottom of the screen. In each trial, in the first display (Screen 1), two object pairs were presented (rings & hats, rings & bags, watches & hats, or watches & bags) either on the left or right side of the screen with objects from the same category shown vertically aligned, and participants were asked to choose either the top or the bottom object set as more expensive (Decision 1). After an intervening blank interval, a second display (Screen 2) was presented. In addition to Screen 1 objects, Screen 2 contained two new object pairs from the remaining object categories, and participants chose the four-object set on the top or bottom as more expensive (Decision 2).

To manipulate the additional resource expenditure, in two-thirds of the trials, the objects shown in Screen 1 were spatially rearranged in Screen 2. The 96 experimental trials were divided into 3 groups of 32 trials and assigned to three layout change conditions: no-swap, within-swap and between-swap. As shown in Figure 1, in the no-swap condition, Screen 1 objects were shown in identical spatial locations in Screen 2. In the within-swap condition, Screen 1 objects on the top or bottom of the display maintained their vertical position in Screen 2 but were horizontally swapped across screens. Finally, in the between-swap condition, Screen 1 objects maintained their horizontal position in Screen 2 but were vertically swapped across screens.

For each participant, objects were randomly assigned to layout change conditions. In addition to the 96 experimental trials, four practice trials were created using objects that were not used in the experimental trials.

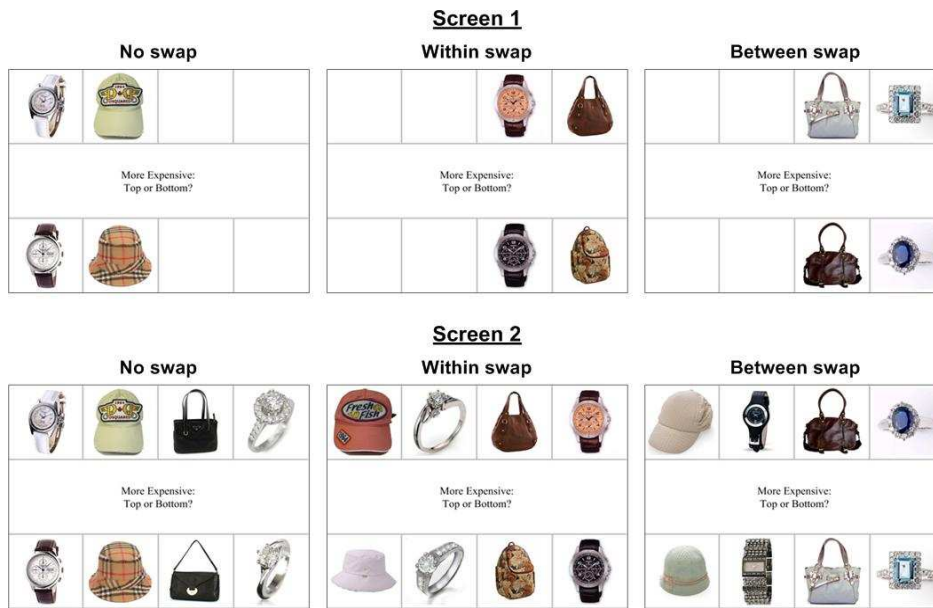


Figure 1. An illustration of the value judgment task and the layout change manipulation (see text for details).

### 4.3 Procedure

Stimulus displays were presented on a 19-in. Viewsonic monitor. The participants' monitor was set to a resolution of 1600 x 1200 and a refresh rate of 85 Hz. The participants were seated 60 cm from the display. They were instructed to choose the more expensive set of objects in both Screen 1 and 2 in each trial and indicate their choice by pressing the corresponding (top or bottom) button on a button box. A participant initiated the trial sequence in each of the 4 practice trials and the subsequent 96 experimental trials by pressing a button on a button box resulting in the presentation of Screen 1. Following the response by participants, the display was blanked for an interval, and then Screen 2 was shown until the participants indicated their final choice.

### 4.4 Measures

Choice and response time for each judgment stage were recorded by the computer as dependent measures. Effort expending is frequently measured by examining time spent completing the task (Bettman et al 1990). Additionally, based on participants' choices concerning objects that were presented in both Screen 1 and 2, we distinguished between decisions that were consistent (i.e., the chosen object set in Decision 1 was part of the chosen object set in Decision 2; Decision 1 = Decision 2) and decisions that were inconsistent (i.e., the chosen object set in Decision 1 was not part of the chosen object set in Decision 2; Decision 1 ≠ Decision 2).

## 5. RESULTS

### 5.1 Choice Consistency Rates

To explore the findings from the present experiment, we began by analyzing consistency rates. In each trial, regardless of the presence or absence of a layout change, the decision sequence was classified as consistent or inconsistent based on whether or not the chosen object pair from Decision 1 was part of the chosen set in Decision 2. That is a decision sequence was defined as consistent when the chosen objects in Decision 1 were

part of the chosen set in Decision 2. In contrast, a decision reversal or inconsistency occurred when the chosen objects in Decision 1 were part of the non-chosen set in Decision 2. The average percentage of consistent trials (consistency rate) was then computed for each layout change condition (no-swap:  $M = 76.02$ ,  $SD = 7.6$ ; within-swap:  $M = 75.71$ ,  $SD = 10.9$ ; between-swap:  $M = 62.10$ ,  $SD = 10.6$ ).

In Hypothesis 1, we expect that the consistency rates will be the same across the no-swap and within-swap conditions. The result supported Hypothesis 1. Consistency rates did not differ across the no-swap and within-swap conditions ( $t < 1$ ) indicating that the within-swap layout change did not impact the extent to which participants' preliminary decision (Decision 1) figured in their final choice (Decision 2). While Hypothesis 2a suggests that the consistency rate will be lower in the between-swap condition than in the no-swap condition, Hypothesis 2b predicts no difference. The result supported Hypothesis 2a. Both the no-swap and within-swap conditions produced somewhat higher consistency rates than the between-swap condition (both  $t_s > 4.58$ , both  $p_s < 0.001$ ).

## 5.2 Response Times

Next we analyzed RTs in Decision 1 and Decision 2 across the layout change by consistency conditions (see Figure 2). In Decision 1, while in the no-swap condition there was no difference in response time (RT) as a function of consistency ( $t < 1$ ), in both the within-swap and between-swap conditions RTs were significantly longer in inconsistent than consistent decision sequences (both  $t_s > 2.12$ , both  $p_s < 0.05$ ). This resulted in a significant layout change by consistency interaction ( $F(2,38) = 4.16$ ,  $p < 0.05$ ). Consistent with Hypothesis 3a, this effect indicates that some aspect of Decision 1 is predictive of the likelihood of a decision reversal in Decision 2. Specifically, a layout change that followed a difficult preliminary decision (i.e., as reflected by longer RTs likely due to a smaller perceived difference between alternatives) was associated with a higher likelihood of a decision reversal or inconsistency, and this finding held regardless of whether or not this layout change occurred within or between alternatives.

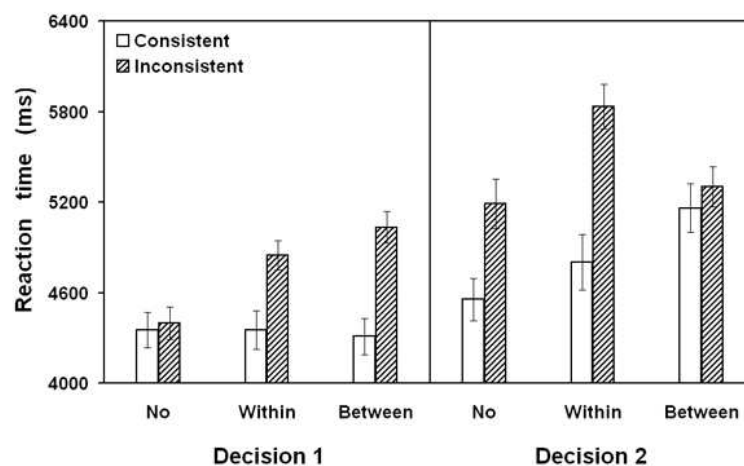


Figure 2. Reaction times for Decision 1 and 2 by consistency and layout change conditions.

In addition, an examination of RTs in Decision 2 revealed that the effects of consistency varied markedly across layout change conditions ( $F(2,38) = 6.50$ ,  $p < 0.01$ ). Specifically, while in the no-swap and within-swap conditions RTs were longer in inconsistent than consistent decisions (both  $t_s > 2.26$ , both  $p_s < 0.05$ ), in the between-swap condition there was no difference in RT as a function of consistency ( $t < 1$ ). The absence of a consistency effect on RT in the latter condition does not imply an absence of processing costs associated with a decision reversal. Rather it is due to longer RTs in consistent trials in the between-swap condition as compared to the other conditions (both  $t_s > 2.80$ , both  $p_s < 0.05$ ). This slowing of RT in consistent trials in the between-swap condition is likely due to the processing costs involved in re-mapping of objects to decision alternatives (i.e., top or bottom) that is required in this condition.

## 6. CONCLUSION

In this study, we investigate the impact of process-induced decision costs of previous decision on subsequent decision. The goal of the present study is to examine the accountability (i.e., “effort-as-information” or “resource availability”) of the impact of previous decision. In the experiment, after the preliminary judgment, the amount of information was controlled but additional resource expending was imposed. Hence, the effect of layout change, if any, can be attributed to the explanation of resource availability. Lower consistency rate occurred when additional resources were required to re-mapping of objects to decision alternatives. The decreased consistency rate implies that resource availability play a significant role in sequential decision-making situations.

Further, the amount of effort spending on preceding decisions also influences subsequent decisions. When the first judgment consumed more resources, the performance of subsequent activities was impaired. Meanwhile, the visual display change raises processing difficulty and impedes fluent processing, which may influence consumer judgments. Again, the data supported that the process of making a difficult preliminary decision (as reflected by longer response times) can deplete self-regulation resources, producing a higher likelihood of a decision inconsistency followed by a layout change. In sum, the expenditure of self-regulation resources impacts not only subsequent performance of cognitive activity but also sequential decision-making results.

This study contributes to the consumer behavior research by investigating when and the extent increasing processing costs (i.e., require more resources) alters the tendency of the subsequent decision to go with the previous decision. Most importantly, the management implication of this study indicates the popular use of dynamic web pages in online shopping situations is likely to increase processing costs by changing product locations which may potentially influence consumer judgments. Both consumers and managers should be aware of such underestimated effects.

## REFERENCES

- Arkes, H. R. and Blumer, C., 1985. The Psychology of Sunk Cost. *Organizational Behavior and Human Decision Processes*, Vol. 35, No. 2, pp. 124-140.
- Baumeister, R. F. and Heatherton, T. F., 1996. Self-Regulation Failure: An Overview. *Psychological Inquiry*, Vol. 7, No. 1, pp. 1-15.
- Bettman, J. R. et al, 1990. A Componential Analysis of Cognitive Effort in Choice. *Organizational Behavior and Human Decision Processes*, Vol. 45, No. 2, pp. 111-139.
- Bettman, J. R. et al, 1998. Constructive Consumer Choice Processes. *Journal of Consumer Research*, Vol. 25, No. 3, pp. 187-217.
- Biyalogorsky, E. et al, 2006. Stuck in the Past: Why Managers Persist with New Product Failures. *Journal of Marketing*, Vol. 70, No. 2, pp. 108-121.
- Brockner, J., 1992. The Escalation of Commitment to a Failing Course of Action: Toward Theoretical Progress. *Academy of Management Review*, Vol. 17, No. 1, pp. 39-61.
- Chen, H. and Rao, A., 2002. Close Encounters of Two Kinds: False Alarms and Dashed Hopes. *Marketing Science*, Vol. 21, No. 2, pp. 178-196.
- Garbarino, E. C. and Edell, J. A., 1997. Cognitive Effort, Affect and Choice. *Journal of Consumer Research*, Vol. 24, No. 2, pp. 147-158.
- Godek, J. et al. 2001. Effort as Information: The Influence of Perceived Decisional Effort on Willingness to Pay. *Proceedings of Society for Judgment and Decision Making*. Orlando, Florida
- Im, H. et al, 2010. The Perceptual Fluency Effect on Pleasurable Online Shopping Experience. *Journal of Research in Interactive Marketing*, Vol. 4, No. 4, pp. 280-295.
- Johnson, E. J. et al, 1988. Information Displays and Preference Reversals. *Organizational Behavior and Human Decision Processes*, Vol. 42, pp. 1-21.
- Kim, J. (2008). *The Influence of Previous Decision on Subsequent Decision*. Unpublished Dissertation, University of Minnesota.
- Klein, N. M. and Yadav, M. S., 1989. Context Effects on Effort and Accuracy in Choice: An Inquiry into Adaptive Decision Making. *Journal of Consumer Research*, Vol. 15, pp. 411-421.

- Kruger, J. et al, 2004. The Effort Heuristics. *Journal of Experimental Social Psychology*, Vol. 40, No. 1, pp. 91-98.
- Loewenstein, G. F., 1996. Out of Control: Visceral Influences on Behavior. *Organizational Behavior and Human Decision Processes*, Vol. 65, No. 3, pp. 272-292.
- Loewenstein, G. F. and Issacharoff, S., 1994. Source Dependence in the Valuation of Objects. *Journal of Behavioral Decision Making*, Vol. 7, No. 3, pp. 157-168.
- Luce, M. F., 1998. Choosing to Avoid: Coping with Negatively Emotion-Laden Consumer Decisions. *Journal of Consumer Research*, Vol. 24, No. 4, pp. 409-433.
- Meyers-Levy, J. and Tybout, A. M., 1997. Context Effects at Encoding and Judgment in Consumption Settings: The Role of Cognitive Resources. *Journal of Consumer Research*, Vol. 24, No. 1, pp. 1-14.
- Monga, A. and Rao, A. R., 2006. Domain-Based Asymmetry in Expectations of the Future. *Organizational Behavior and Human Decision Processes*, Vol. 100, No. 1, pp. 35-46.
- Novemsky, N. et al, 2007. Preference Fluency in Choice. *Journal of Consumer Psychology*, Vol. 44, No. 3, pp. 347-356.
- Samuelson, W. and Zeckhauser, R., 1988. Status Quo Bias in Decision Making. *Journal of Risk and Uncertainty*, Vol. 1, No. 1, pp. 7-59.
- Schmeichel, B. J. et al, 2003. Intellectual Performance and Ego Depletion: Role of the Self in Logical Reasoning and Other Information Processing. *Journal of Personality and Social Psychology*, Vol. 85, No. 1, pp. 33-46.
- Schwarz, N. and Clore, G. L., 2006. *Feelings and Phenomenal Experiences*. In Higgins, E. T. and Kruglanski, A. (Eds.), *Social Psychology. A Handbook of Basic Principles* (2nd ed.). Guilford Press, New York.
- Shugan, S. M., 1980. The Cost of Thinking. *Journal of Consumer Research*, Vol. 7, No. 2, pp. 99-111.
- Song, H. and Schwarz, N., 2008. If It's Hard to Read, It's Hard to Do: Processing Fluency Affects Effort Prediction and Motivation. *Psychological Science*, Vol. 19, No. 10, pp. 986-988.
- Staw, B. M., 1981. The Escalation of Commitment to a Course of Action. *Academy of Management Review*, Vol. 6, No. 4, pp. 577-587.
- Thaler, R. H. and Johnson, E. J., 1990. Gambling with the House Money and Trying to Break Even: The Effects of Prior Outcomes on Risky Choice. *Management Science*, Vol. 36, No. 6, pp. 643-661.
- Vohs, K. and Baumeister, R. F., 2004. *Understanding Self-Regulation: An Introduction*. In Baumeister, R. F. and Vohs, K. (Eds.), *Handbook of Self-Regulation: Research, Theory, and Applications* (pp. 1-12). The Guildford Press, New York.
- Vohs, K. et al, 2008. Making Choices Impairs Subsequent Self-Control: A Limited Resource Account of Decision Making, Self-Regulation, and Active Initiative. *Journal of Personality and Social Psychology*, Vol. 94, No. 5, pp. 883 - 898.
- Vohs, K. and Schmeichel, B. J., 2003. Self-Regulation and the Extended Now: Controlling the Self Alters the Subjective Experience of Time. *Journal of Personality and Social Psychology*, Vol. 85, No. 2, pp. 217-230.
- von Helversen, B. et al, 2008. Exploring the Hardship of Ease: Subjective and Objective Effort in the Ease-of-Processing Paradigm. *Motivation and Emotion*, Vol. 32, No. 1, pp. 1-10.





# Short Papers



# TESTING NEW SCENARIOS FOR DISTANCE LEARNING, BASED ON THE INTRODUCTION OF NEW DEVICES

Victor Ximenes and Gabriel Pontual

*C.E.S.A.R - Recife Center for Advanced Studies and Systems  
Rua Bione, 220 Cais do Apolo, Recife-PE, Brasil.  
CEP: 50030-390*

## ABSTRACT

Distance learning is, nowadays, an important modality of higher education in Brazil. Since it employs a diversity of technological resources, it has presented a substantial development lately. However, a significant number of web-based courses have unwittingly or naively endorsed a transmission model of learning similar to what happens in a traditional, face-to-face classroom, which impoverishes the distance learning experience. Introducing user centered design methods to investigate distance learning experience with the adoption of new technological devices allowed us to explore interaction scenarios and create design recommendations from tests with distance learning students. The purpose of this paper is to briefly describe the first iteration of research and user testing of prototyped interactions in distance learning, conducted by the usability experience design team at C.E.S.A.R (Recife Center for Advanced Studies and Systems).

## KEYWORDS

Distance Learning, Usability Testing, Experience Prototyping, User Centered Design.

## 1. INTRODUCTION

Distance learning is a modality of education where students do not need to be physically collocated, as in face-to-face classrooms. Another essential (but not general) characteristic is that students are allowed to create their learning schedules as the class becomes asynchronous and immaterial. Students and teacher get connected through different medias, such as radio, CD's, letters, books, television programs and recently, hypermedia such as mobile platforms and Internet.

Distance learning, though, has not always had its present purpose. According to Azevedo (2000), in Brazil it used to be considered an inferior education modality. When other modalities failed to supply minimum quality education to the population, distance learning used to play the role of an emergency and partially effective teaching system. Such palliative aspect of distance learning contributed to make it socially known as a modality of education for low-income classes. The turning point, however, occurred when new interactions became possible with the insertion of innovative technologies such as Internet and a number of tangible interfaces. It all brought about a number of benefits to learning such as construction of collaborative learning online communities (Azevedo, 2000), accessibility and increased sensory engagement (Pontual Falcão, 2007). Students have the chance to transform classes in a much more interactive and collaborative situation. Nowadays distance learning is a viable and desirable option for learning and represents an important modality of higher education in Brazil. According to the 2009 College Education Census from the Brazilian Ministry of Education, distance learning has presented a 98.9% growth rate of enrollments between 2007 and 2008, and distance learning students represent 14.3% of higher education students in Brazil (Ministério da Educação, 2009).

In the second semester of 2010, a multidisciplinary team of user experience designers from C.E.S.A.R (Recife Center for Advanced Studies and Systems), started a study on distance learning education based on User Centered Design Process, comprehending the following phases: Research, Ideation, Prototyping and Testing. In the research phase, interviews and a literature review were conducted focusing on the Brazilian distance learning context (higher education). Next, at the Ideation phase, new possible and plausible interaction scenarios supported by the introduction of new technological devices were brainstormed. A few

possible scenarios (feasible to be implemented in the near future) had their interaction prototyped and evaluated in laboratory at the Prototyping phase. Finally, user tests aimed to observe, evaluate and understand how current distance learning students would react to the proposed scenarios, which characteristics from the proposed system should be adapted and how students would behave on a virtual live class situation (where 'live class' is understood as a synchronous moment, supported by internet streaming) designed to be installed in their living rooms. Introducing user centered design methods to investigate distance learning experience with the adoption of new technological devices allowed us to explore interaction scenarios and create design recommendations by testing them with real distance learning students. The purpose of this paper is to briefly describe the first iteration of research and user testing of prototyped interactions in distance learning as well as its outcomes.

## 2. INTERACTION SCENARIOS

After interviewing a number of Brazilian distance learning college teachers and students, and also conducting a literature review, it was possible to understand that their chief difficulty is not technological, but conceptual. A significant number of web-based courses have unwittingly or naively endorsed a transmission model of learning similar to the traditional face-to-face classroom in their designs. Therefore, if the web is to be used as a versatile medium for learning, then, careful attention must be paid to the design of interaction that can foster the negotiation of meanings and the construction of knowledge through social interactions. Vicarious interaction within the class as a whole may result in greater learner satisfaction than will the overt engagement of each participant.

According to the research, traditional classroom lessons generally include different moments of interaction: (1) Exposition, when the teacher explicates the content, (2) Problematizing, when teacher and students enter a question-answer-feedback mode, (3) Abbreviation, students solve problems on the knowledge domain. A scenario was designed to take advantage of the ever increasing number of technological devices such as tablets and docks, making use of them in order to facilitate the conceptual paradigm shift from an inappropriate face-to-face adapted model to a stage where technology would improve social interaction by acting on the classroom lessons moments previously mentioned.

Based on research, a number of technology-assisted interaction scenarios were built to support different virtual live class moments. They were divided into two groups: (i) possible, the ones that are feasible to be implemented in the near future and with few enhancements on the technology devices currently available in the market, (ii) plausible, the scenarios whose implementation is strongly dependent on the future evolution of technology. For practical reasons, this research focused on the possible scenarios.

A possible scenario was designed to be a tangible interface for distance learning, allowing interaction through a tablet (such as the Apple's iPad) synchronized wirelessly with a dock (a device that is able to receive and synchronize signals between the tablet and TV) connected to internet and to a television placed in students houses. In this scenario the prototyped dock was equipped with fingerprint reader, camera and microphone (recognizing voice commands, fingerprints, faces and gesture commands), and touch screen tablet with fingerprint reader, camera and microphone. The three devices (tablet, dock and TV) are meant to be connected through a distance learning ambiance-simulated software.

It is important to mention that the tested scenarios contemplated a simulated distance learning virtual live class with both individual and a group of students. In the individual student scenario, each student should be located in his or her house with the described infrastructure and equipment. In the group scenario, up to three students shared the same tablet during the class. The latter scenario has been built according to distance learning research findings, which revealed that in Brazil students often share the same infrastructure. The expensive access to technology devices has also been considered, as according to Kunze (2011), Brazilian tax rates on imported electronics increase the price of electronic devices to the point that the most expensive iPad in the world, for example, is sold in Brazil.

A couple of virtual live class moments (based on exposition and problematizing interaction moments) were then chosen to support the aforementioned scenario and had their tasks prototyped in the laboratory tests: Presence, when students are recognized by their devices and are clustered on the class user interface based on some pre-defined criteria; Question, when the teacher poses a multiple choice question to all students and receives the answer.

### 3. EVALUATION METHODS

The user testing lasted three days and was divided in two stages: sessions with individuals and groups of users. The C.E.S.A.R's process of user testing was adopted, which involves the following phases: Planning, Recruiting/Task Design, Preparation, Execution and Analysis (Ximenes, 2008).

For the preparation phase, the demanded infrastructure involved a usability test laboratory equipped with cameras, a couch and a television, a generic dock for MP3 players and a handheld pocket PC (to simulate devices), and low fidelity prototypes of interactive presentations to simulate software. According to test tasks, two techniques have been adopted: semi-structured interviews and the "Wonderful Wizard of Oz" technique (Buxton, 2007).

#### 3.1 Semi-structured Interviews

According to Denscombe (2007), in a semi-structured interview, the interviewer has many questions to be answered but is prepared to be flexible in terms of the order in which the topics are considered, and to allow interviewees to develop ideas and speak more widely on the issues raised by the researcher.

In the present research, the recruited participants, after being introduced to test scope, were invited to talk openly about their routines as distance learning students, which course they attended, etc. Other interview topics included describing details of how their classes and online virtual learning environment are currently structured, how they use this structure on a daily basis, and their familiarity with technology and devices. Many interviewees were not only students but also employees in their respective colleges, providing the research with rich information about distance learning institutions.

#### 3.2 The "Wonderful Wizard of Oz" Technique

"Wonderful Wizard of Oz" is a user experience sketch technique proposed by Buxton (2007). It proposes the employment of low fidelity prototyping with high fidelity user experience instead of high fidelity prototyping and mediocre experience. In other words, while aesthetics are important in the design of a product or interface, they are far less valuable for a user's experience than are the functions and the available cues it offers.

In the present work, this specific technique has been employed during tests to simulate the software features and to evaluate the experience on how users would execute three specific tasks: (1) login into the system, (2) watch a teacher exposition and (3) answer a multiple choice question. Software sketches were constructed with simple and interactive PowerPoint presentations, where buttons simulated the interactivity.

In the first task (login), both tablet and TV showed a welcome screen inviting users to login through one of four options: fingerprint recognition, voice recognition, face recognition (available both on tablet or dock) and typing user and password (tablet only). Evaluation of this task aimed to find out which of the four alternatives was the preferred form of login and in which device (tablet or dock) user would input data to login. In the second task, an 18 minutes video lecture was presented. This task aimed at observing how users behave (ergonomic and postural aspects, attention levels, etc.) while attending to a distance learning live class. Finally, in the third task, participants were asked to inform the correct answer to the system using the interactional option most appropriate to them (voice commands, hand gestures or touching the device screen). For the group tests, the participants were encouraged to discuss with each other before informing the correct response.

### 4. RESULTS

The results from the user test sessions were analyzed qualitatively. It is important to reinforce that qualitative methods are not appropriate to determinate averages since such methods do not comprehend a significative sample to be statistically relevant. The present analysis aimed at identifying patterns of user's behaviors and attitudes (as well as their comments) in comparison with all data collected during the research phase

(interviews and literature review). Finally, a list of design recommendations was created regarding the interaction scenarios and live class moments:

- When invited to explore technology, users preferred to use biometric recognition on the tablet as the way to log in to the system;
- When invited to interact with voice commands users preferred to dialog with the equipment instead of just executing short word commands. It was observed that in many occasions, users established short dialogs with the machine. Apparently, talking to the equipment in ways we can currently handle requires more familiarity with technology, since individuals are not used to this form of interaction;
- In task 2 the tablet faded out from users' actions and attention. If we are to raise users' levels of awareness to this equipment during classes, making it more interactive, it is recommended to further study possible tasks and forms of interaction based on vibration and sound, for instance;
- When invited to respond to a multiple-choice task, users preferred to touch the correct alternative on the tablet's screen;
- Voice recognition seems important, but users demonstrated preference for a conversation channel rather than a single answer channel.

## 5. CONCLUSION

Many distance learning systems were found to inadequately transfer parameters from face-to-face learning to virtual environments. In order to experiment innovative approaches for distance learning systems, prototyped interactions and user tests were designed to introduce new technological devices such as tablets and docks in distance learning classes. However, one of the biggest challenges was to engage the participants during the test sessions. Since they are not familiar with this use of technology in their daily activities, sometimes it was difficult to make them actually feel as a part of the tested scenarios.

The perspectives for this topic are promising since technology is continually evolving with brand new interaction forms that go beyond the graphical user interface. Unfortunately, further data concerning the results of this work-in-progress cannot be revealed due to non-disclosure agreement. As for future work, there is still a number of scenarios to be investigated as well as live class moments to be explored, replicating the same experience prototyping and testing methods. At this stage, particular attention must be paid to the development of a prototype based on the design recommendations elicited by the user experience team. It is important that this prototype is iteratively tested with users and incrementally refined.

As described in this paper, using the experience prototyping technique to investigate and explore new scenarios of interactions in a user centered design process allowed us to discover very early which strategic decisions should be taken regarding technology and user interface. Because this research has been conducted since the early stages of the project and the engineering team was able to work along with the user experience team, software and hardware recommendations could be made more consistently and quickly. All the described techniques and methods have proved to be valuable in C.E.S.A.R and were successfully applied in projects from contexts other than distance learning, such as the development of healthcare products, banking automation and taxicab computerization devices.

## ACKNOWLEDGEMENTS

Authors are thankful to Taciana Pontual Falcão and Georgia Barbosa for their support during the production of this paper.

## REFERENCES

- Bia Kunze, 2011. Por que tudo no Brasil é tão caro? *Tecnoblog*, [blog] 3 January. Available at: <<http://tecnoblog.net/52271/custo-brasil-por-que-tudo-no-brasil-e-ao-caro>> [Accessed 22 February 2011].
- Buxton, B., 2007. *Sketching User Experiences*. 3rd ed. San Francisco: Morgan Kaufmann Denscombe, M., 2007. *The Good Research Guide for small-scale social research projects*. 3rd ed. New York: McGraw Hill.
- Ministério da Educação do Brasil, 2009. *Censo INEP 2009 para o Ensino Superior*. [online] Available at: <<http://www.inep.gov.br/superior/censosuperior/default.asp>> [Accessed 25 February 2011]. Pontual Falcão, T. and Gomes, A.S., 2007. Tangible Interfaces for Education. *Proceedings of 3rd SBIE - Brazilian Symposium of Computer Education*. São Paulo, Brasil, pp. 2. Wilson Azevedo, 2000. *Panorama Atual da Educação a Distância no Brasil*. [online] Available at: <[http://www.revistaconecta.com/conectados/wilson\\_seminario.htm](http://www.revistaconecta.com/conectados/wilson_seminario.htm)> [Accessed 25 February 2011].
- Ximenes, Victor and Levi, F., 2008. Integrando testes de usabilidade em projetos sigilosos. *IHC '08 Proceedings of the VIII Brazilian Symposium on Human Factors in Computing Systems*. Porto Alegre, Brasil, pp. 304-307.



# USING VITAL-SENSOR IN TRACKING USER EMOTION AS A CONTEXTUAL INPUT FOR MUSIC RECOMMENDATION SYSTEM

Nguyen Thuy Le<sup>1</sup>, Jin Nakazawa<sup>2</sup>, Kazunori Takashio<sup>2</sup> and Hideyuki Tokuda<sup>2</sup>  
<sup>1</sup>*Faculty of Environment and Information Studies, Keio University*  
<sup>2</sup>*Graduate School of Media and Governance, Keio University*

## ABSTRACT

Emotion is a novel contextual key for Music Recommendation System to help users to manage their expanding digital music libraries. In this paper, we propose EmuPlayer - a Music Recommendation System (MRS) which tracks a user's emotion and suggests songs in the form of a playlist that is sorted to match the user's current emotion. For a particular emotional state of the user, the system evaluates songs according to two factors: the user's preference, and the potential of a song to influence on the user's emotional state. We evaluated both EmuPlayer Emotion Recognition accuracy and its efficiency in Recommending songs.

## KEYWORDS

Music Recommendation, Emotion Recognition, Vital-sensor, Emotional Model.

## 1. INTRODUCTION

As there is a close relationship between music and the emotional state of its listeners, users often choose songs based on their current feelings. However, the constantly expanding storage of music devices makes it difficult for them to do so. Though there have been a remarkable number of studies on MRS, a MRS which handles emotional state of a user as real-time input has not been proposed. That is the motivation behind this research, with two goals which are (1) To suggest relevant songs to the user's current emotion; and (2) To avoid songs which are potentially harmful to the user's emotional state. Major features of EmuPlayer are the following four fold: (1) to track User Emotion; (2) to give recommendation playlist by sorting songs based on the user's current emotion; (3) to evaluate the effect of played song; and (4) to rate songs by two factors: the relevance to user preference, and the effect on user emotional state. The work of evaluating and giving feedback about how a song influences on the user's emotional state after he listened to it is highly desired, especially in case a song gives bad effect. Therefore, the first and the fourth features are considered conceptually distinct point of EmuPlayer.

## 2. EMOTION RECOGNITION USING VITAL-SENSOR AND RUSSELL MODEL

Among available methods for Emotion Recognizing, such as employing brainwave, gesture, eyes movement, facial expression, method using vital-sensor appears to be mostly applicable to EmuPlayer. There are evidences proving that emotion can give rise to effective experiences such as feelings of arousal or pleasure [9] [11], and arousal and pleasure express themselves very clearly through vital information. Therefore, by using vital-sensor, we can assess levels of arousal and pleasure, and deduce current emotional state of the subject. In addition, Vital-sensor transcends in rapidly decreased size, the continuity and the sensitiveness of the output. Above factors reason the work of employing vital-sensor, RF-ECG, to EmuPlayer.

To map emotional states from vital information, we use Russell Model, which is a dimensional model presenting how vital information and emotions are related. In Russell Model, 8 emotions are clearly presented with exact angle of  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ ,  $135^\circ$ ,  $180^\circ$ ,  $225^\circ$ ,  $270^\circ$ ,  $315^\circ$ , on the two dimensional space where horizontal axis is Pleasure and Vertical axis is Arousal [10]. Skin Temperature was chosen to determine the level of Pleasure as it increases towards pleasure and decreases on the contrary [7], and Heart Rate to determine the Arousal state as high arousing state results in fast heart rate, and low arousing state results in slow heart rate [5][2]. In order to assure the system's portability, we employed as few biosignal as possible. RF-ECG was chosen for the same reason as it can measure both heart rate and skin temperature.

According to the theory of the Fuzziness of words, words which are close together in the circular ordering have considerably the same meaning [10]. This affirmation provides foundation to respectively divide Russell model into eight equal parts representing the 8 emotions.

Mapping then is done by rounding the angle defined by user's reading to the closest emotion region as shown in Figure 1. While the origin position represents the user's normal values of skin temperature and heart rate, the mapping scale value is set as the variance of corresponding heart rate and skin temperature measured at the user's normal state.

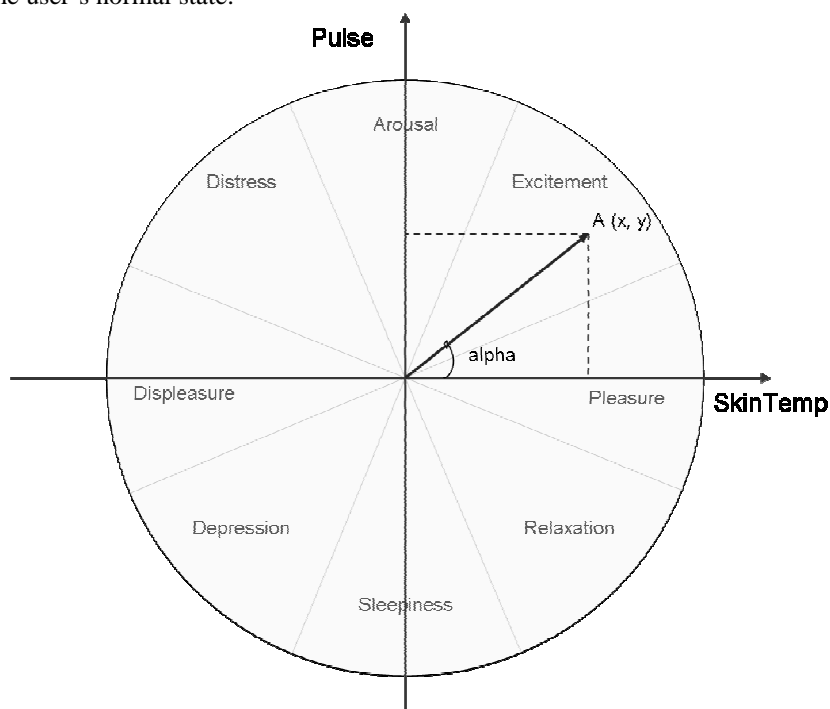


Figure 1. Mapping user emotion representing point onto Russell model

### 3. MUSIC RECOMMENDATION

As EmuPlayer evaluates songs regarding to 2 factors as discussed above, there are two subjects to solve: (1) to study User Preference, (2) to study the potential of songs to influence on emotional states of users.

**Study on User Preference and Songs Emotional Effect:** The concept used to study the user's preference is familiar and has been applied in many systems before. The user will rate like or dislike to songs after listening. Then, all users' listening history along with their preference will be recorded and stored inside the database for later reference. The potential of a song to influence on user emotion can be decided by comparing emotions before and after that user listened to the song. Russell Model is separated into two regions: (1) include bad emotions (Distress, Displeasure, Depression), and (2) include the left ones. Movements from (1) to (2) are Good effect. As emotions belonging to region (1) get worse by the order of Distress, Displeasure and Depress, according to Russell's concept and their position on the model, within (1),

movements in which former emotion is better than the later one represent Good effect, and vice versa. Inside (2), for that definition of which emotion is better than the other is different from subjects, and for that all those emotions are not harmful to emotional states of users, movements within (2) are assessed as bringing no harm and therefore stated as Normal.

**Song Rating:** Songs are rated by the following scale: like (+1), dislike (-1), good influence (+1) or bad influence (-1). The work of rating helps to assure that songs with higher quality always rank higher in the playlist. So, by recommending users to pick up high ranking songs, users are expected to have the best song choice which matches their current emotion. Scoring equation is as follows.  $Score = ((current\_score * listened\_time) + new\_score) / listened\_time + 1$ .

## 4. EVALUATION

### 4.1 Evaluation on Emotion Recognition

**First experiment** was carried to test the accuracy of Emotion Recognition. Participants were driven to particular feelings through arranged situations. At the same time, the engine detected their feelings. In order to check whether participants experienced the experimenting emotions or not, after each situation, participants were asked to do survey with two questions: (1) if they experienced the emotion expressed through the situation, and (2) if not, what emotions rather than the one in (1) they experienced.

Table 1 shows experiment's result. Input are emotions tested through arranged situations, verified by participants. Output is emotions detected by the engine. Matching pairs of input/output is marked as 1. Matching pairs between input and emotions which are expressively close to input marked as 2 show that most of misdetected parts fall into closest emotions of the experiment ones. The overall accuracy is 64.5%.

**Second experiment** was carried in order to (1) verify whether the system can realize user's emotional changes, and (2) verify songs' influence on listeners' emotions. In this experiment, we let participants listen to music and observe their emotional states' changes. Four cases were performed (see Table 2). The result shows that the system responses precisely to the changes of participants' emotion. In each of following experiments, about 10 to 12 participants took place. They were all male. Average age is 21 years old.

Table 1. Emotion Recognition First Experiment

Output (Result from Engine) \ Input (Verified Experimenting Emg)	Arousal	Excitement	Pleasure	Relaxation	Sleepiness	Displeasure	Distress	Depression
Relaxation	4.25%	7.96%	22.77% [2]	55.58% [1]	9.44% [2]	0	0	0
Excitement	10% [2]	63.34% [1]	21.66% [2]	5%	0	0	0	0
Pleasure	1.4%	10.87% [2]	66.86% [1]	20.87% [2]	0	0	0	0
Arousal	81.68% [1]	5% [2]	0	6.66%	6.66%	0	0	0
Depression	8.33%	0	0	0	6.66%	20%	10% [2]	55.01% [1]

Table 2. Emotion Recognition Second Experiment

Case	Experiment	Mean	Result
1	Arousal → Pleasure/Relaxation	Classical music	Arousal: 81.68% → 1.41% Pleasure: 0% → 54.91% Relaxation: 6.66% → 38.94%
2	Normal → Pleasure	Music participants like	Pleasure: 66.86% → 93.33%
3	Normal → Excitement	Fast beat Music	Excitement: 10.87% → 62.12%
4	Normal → Depression	Loud Heavy Music played in long time	Pleasure: 66.86% → Depression: 80.02%

**Conclusion 1:** (1) Accuracy of extracting emotion is 64.5%; (2) The system is strong at detecting bad emotions; (3) The system detects precisely regarding to changes in user emotion; and (4) Hypothesis of music influencing on user emotion is true.

### 4.2 Music Recommendation Efficiency

In order to evaluate the performance of EmuPlayer, records of users' listening history was studied to observe firstly the status of high-rating songs as shown in Figure 3. Due to the proposal, high-rating songs mean good recommendation for a user by his current emotion; hence the work of observing high-rating songs' status is

important in evaluating the quality of recommending. Besides, how emotion changes in cases where songs were “liked” by the users will also be observed as shown in Table 4. Following is the result.

Table 3. Status of high rating songs

State of high-score songs after users listened to them	Percentage
Disliked	6.66%
Bad effect on user emotion	0%
Reduced in score	8.57%

Reasons of being reduced in score	Percentage
Disliked	6.66%
No longer “Liked”	1.91%
Bad effect on user emotion	0%

Table 4. Emotion changes in “liked” song

Emo Before	Emo After	Percentage
Arousal	Relaxation	1%
Arousal	Sleepiness	1%
Excitement	Arousal	1%
Excitement	Pleasure	2%
Pleasure	Excitement	2%
Pleasure	Pleasure	77%
Relaxation	Arousal	1%
Relaxation	Pleasure	2%
Relaxation	Relaxation	7%
Sleepiness	Sleepiness	4%
Displeasure	Displeasure	2%

Percentage of high-rating songs disliked by users is 6.66%. Percentage of songs paid bad effect on users’ emotion is 0%. And percentage of high-rating songs reduced in score is 8.57%. In this 8.57% songs, number of songs reduced in score because of being disliked is 6.66%, no longer liked by users is 1.91%, and causing bad effect on users’ emotion is 0%. Overall, 0% of recommending songs caused bad influence on users’ emotion. Besides, according to the count on emotional movements in “liked” songs’ cases shows that, although there was no recommendation giving good effect on users’ emotion referring to the definition of good effect proposed as above, there was no recommendation causing bad emotion effect on users either. Therefore, the system has achieved its goal of avoiding recommendations which are potentially harmful to users’ emotion.

**Conclusion 2:** From all of these observations, another conclusion can be deduced that EmuPlayer algorithm ensures recommendation of songs meeting the proposed two requirements, as songs influencing badly on user emotion holds 0%, and songs being “dislike” in later listening time holds only 6.66%.

## 5. CONCLUSION AND FUTURE WORK

For all the work and experiments proposed so far, it can be concluded that: (1) Concept of EmuPlayer which is: Evaluating song through 2 factors, and Employing User Emotion as crucial input for MRS is essential; (2) Although Accuracy of extracting emotion is not very high: 64.5%, EmuPlayer is Strong at detecting bad emotion which make it reasonable to be applied on giving alert when playing music influences badly on listener’s emotional state purpose; and finally (3) EmuPlayer’s efficiency in suggesting songs meeting the two requirements has been proved.

Besides, in order to enhance the performance of EmuPlayer system to create a stronger Music Recommendation System, following subjects will be admitted as future works, focusing on three main issues which are: (1) to enhance the accuracy of detecting emotion by employing other means than Heart Rate and Skin Temperature and alternating RF-ECG sensor; (2) to enhance the work of Recommending music by combining proposed method with songs’ content analyzing; and (3) to enhance reasoning users’ state or condition by combining User Emotion with context analyzing.

## REFERENCES

- [2] J.T. Cacioppo, L.G. Tassinary, and G.G. Berntson. *Handbook of psychophysiology*. Cambridge University Press, 2007.
- [3] Man-Kwan Shan<sup>2</sup> Fang-Fei Kou<sup>1</sup>, Meng-Fen Chiang<sup>2</sup> and Suh-Yin Lee. Emotion-based music recommendation by association discovery from film music. *ACM*, 2005.
- [4] Gerals L. Clore Justin Storbeck. Affective arousal as information: How affective arousal influences judgments, learning, and memory. *Social and Personality Psychology Compass*, 2:1824-1843, 2008.

- [5] Owen Craigie Mayers. A mood-based music classification and exploration system, 2007.
- [7] Anne J. Blood and Robert J. Zatorre. Intensely pleasurable responses to music correlate with activity in brain regions implicated with reward and emotion. 10.1073/pnas. 191355898 PNAS September 25 2001, vol.98, no.20.
- [9] James A. Russell. A circumplex model of affect. *Personality and Social Psychology*, 39(6):1161-1178,1980.
- [10] James A. Russell. Core affect and the psychological construction of emotion. *Psychological Review*, 110(1):145-172, 2003.

# CULTURA: A TWO CULTURAL MODEL FOR UNDERSTANDING HUMAN BEHAVIOR ON THE INTERNET

Santosh Kumar Kalwar

*Lappeenranta University of Technology  
Lappeenranta, Finland*

## ABSTRACT

This paper addresses the questions of impact of culture on the Internet. It examines seven simple questions such as a) who are you: an Asian/Western? b) what is culture to you? c) how can culture aid in understanding human behaviour? d) what are the cultural issues that affects your feelings on the Internet? e) how do you feel for personalized or customizable webpage? and f) how would you rate five user's cultural characteristics: i) Intuition ii) Privacy iii) Networking iv) Safety v) Self. The literature study showed that none of the followings --Trompenaars, Hampden-Turner and Hofstader's cultural models is suitable for comparing cultures, especially in case of Human-Computer Interaction. Ten university participants from two different culture participated, and the summary of the findings has been presented. Moreover, a two-culture model based on five parameters has been proposed and tested among participants. In order to understand the definition of culture, we asked participants about culture on the Internet. Thus, formulating various definitions of culture based on their feedback.

## KEYWORDS

Human, Anxiety, Internet, Culture, Behavior

## 1. INTRODUCTION

*“Every age, every culture, every custom and tradition has its own character, its own weakness and its own strength, its beauties and cruelties; it accepts certain sufferings as matters of course, puts up patiently with certain evils. Human life is reduced to real suffering, to hell, only when two ages, two cultures and religions overlap.”* ---Hermann Hesse

As pointed out by Hesse in the above lines, there is strength and weakness of each culture during their times and a culture might be susceptible with another culture. It is understood that human civilization has been divided based on country, culture, education and ethnicity. These categories of division have their own benefits and difficulties. A similar idea is still persistent in today's digital age. Recently, it has been pointed out by President Barack Obama that we live in the “age of Internet” (BARACK OBAMA, 2010). In the age of the Internet, culture has greater role to play in building applications, services and culturally driven systems. It seems that technology changes with time whereas a user hardly changes their behavior for the sake of the changing technologies. In short, it is the technology, which has to change, not the users. Alternatively also called, “cultural frame shifting” (Markussen & Krogh, 2008). On one hand, the way we behave on the Internet maybe independent of the places, culture, country, or region. On the other hand, our behavior is seemingly based on well-conceived actions. Based on that, the technological tools that are used in our daily pursuit should consider cultural aspects for design and development. Although measuring value of culture (Snowball, Jeanette D, 2008) could have different meaning based on different context, especially in Human-Computer Interaction (HCI) culture is fundamental in design and development of any system or product. Therefore, based on the rationale of our study, we have proposed following research questions and hypothesis. RQ1: What is the cultural issue affecting human behavior on the Internet? H1: There could be various cultural issues based on anthropological, social, cognitive and psychological nature of an individual or object that affects human behavior on the Internet.

## 2. LITERATURE STUDY

From use and usefulness view of human-computer interfaces, cultural issues of a user's acceptance have been discussed (Davis, 1993; Hubona, 1996). Furthermore, from implementation view, a term "internationalisation - localisation strategy" has been explained (Uren et al., 1995). In literature, the huge data collected by researchers such as Trompenaars and Hampden-Turner proposed seven dimensions of cultures, which are minutely used in HCI (Trompenaars, F. & Turner, C. H. 1997). Hofstede's cultural dimensions are very popular which are quite similar to Trompenaars and Hampden-Turner theory. Prof. Hofstede's dimensions are based on larger sample, which were collected during 1960-1980's. His dimensions include:

**1. Power-distance:** It defined how people in the particular culture accept the unequal power distribution among themselves. It has been categorized into two dimensions: High power distance and Low power distance. High power distance is based on centralized decision-making system whereas lower power distance is based on distributed decision-making system—where everyone expects to make and share decision among him or her.

**2. Collectivism vs. Individualism:** It is defined based on how people in that culture define them to be part of group or themselves.

**3. Femininity vs. Masculinity:** It is defined based on how people in that culture support feminine or masculine attributes.

**4. Uncertainty Avoidance:** It is defined as an extent to which the culture in that society fears and avoids uncertainty.

**5. Long vs. Short-term Time Orientation:** It is defined based on how people in that culture envision themselves.

The model proposed by Hofstede has been used extensively and has been cited widely. However, it has also been criticised for its lack of depth and richness (Baskerville, Ratner, C. & Hui, L. 2003). Hofstede's model also does not consider environmental issues. In fact, there is very limited evidence of environmental issues taken into account for cultural studies in the literature. The environment issues could either technical or non-technical. Especially among HCI, the impact of environment could affect user's behavior.

### 2.1 Culture in HCI: An Example

Ten years ago, Marcus and Gould proposed a design consideration for two websites from two different countries. In their first example, they showed the University website of Malaysia where the country has high power-distance dimension based on Hofstede's model. In their second example, they showed the high school website of Netherlands where low power-distance is demonstrated while considering design guidelines. The two websites vary in feature, focus, and design. The Malaysian website consists of "strong axial symmetry, a focus on the official seal of the university, photographs of faculty or administration leaders conferring degrees, and monumental buildings in which people play a small role. A top-level menu selection provides a detailed explanation of the symbolism of the official seal and information about the leaders of the university" whereas the Dutch website has "stronger use of asymmetric layout, and photos of both genders in illustrations. These websites emphasize the power of students as consumers and equals. Students even have the opportunity to operate a Webcam and take their own tour of the Ichthus Hogeschool" (Marcus, A. & Gould, E. 19 June 2000).

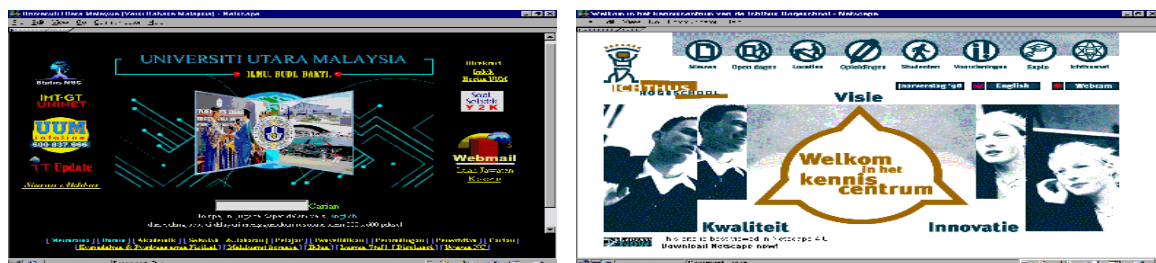


Figure 1. High power distance: Malaysian University Web site and Low power distance: The Dutch high school website (Marcus, A. & Gould, E. 19 June 2000).

### 3. PROPOSED: TWO CULTURE MODEL

We have sincerely rejected the Hofstede's model of cultural dimension since it lacks clarity and richness of user's experience. Thus, a proposed two cultural model based on Environment (E) and user experience has been devised. Majority of the times, users do not change their environment neither do they change their behavior but they "adapt" to the given environment. In case of the technological change, the latter statement may not be feasible. Following attributes based on two different cultures (C and D) has been assumed. Therefore, the model is based on following defined characteristics of user experiences. These characteristics were formulated based on user's feedback and rigorous discussion. The 'two cultural model' is based on set of assumptions, which has been shown below. It has been assumed that if a user belongs to one culture (C) with five mentioned characteristics individually then a user may not be equal to another culture (D). Furthermore, if a user "adapts" to any given set of preferences or characteristics and belongs to an environment (E) then a user might have similarities with another user (B) from different culture (D) which is semantically shown hereunder:

1. **Intuition:** An act of knowing or sensing
2. **Privacy:** A general concern on the Internet
3. **Networking:** The act by which users wants/desires to collaborate
4. **Safety:** A method or technique developed by a user to avoid uncertainties
5. **Self:** A desire to perform better, learn and live

If A is user(s) from a culture (C) with following characteristics:

$A = \{\text{Intuition} \rightarrow \text{Privacy} \rightarrow \text{Networking} \rightarrow \text{Safety} \rightarrow \text{Self}\}$

And B is another user(s) from yet another culture (D) with similar characteristics but with different preferences then

$B = \{\text{Self} \rightarrow \text{Safety} \rightarrow \text{Networking} \rightarrow \text{Privacy} \rightarrow \text{Intuition}\}$

$A \neq B$  if  $A = \{\text{Culture (C)}\}$  and  $B = \{\text{Culture (D)}\}$

$A = B$  iff  $A = \{\text{Culture (C), Culture (D)}\}$  and  $B = \{\text{Culture (C), Culture (D)}\}$

$A=B$  and  $C=D$  if  $(A, B) = \{\text{Environment (E)}\}$

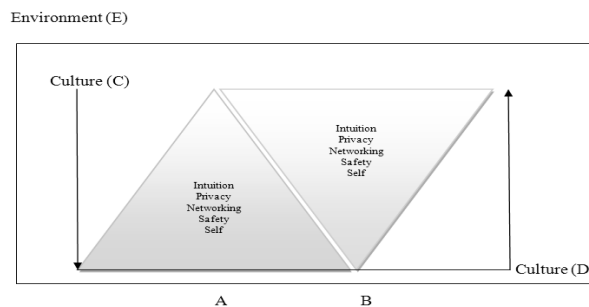


Figure 2. A model for understanding two different cultures (C, and D) by two different users/system (A, and B) on the Environment (E).

### 4. METHODS & RESULTS

We conducted a pilot study with ten university participants. We divided these participants between two cultural groups (C and D). Each group consists of five participants. The first group consisted of Master of Science students from (*xyz*) university of technology studying at Department of Information Technology in (*A*) country. Whereas, the second group consisted of Master of Science students from (*pqr*) university of technology studying at Department of Business Administration in (*B*) country. The first group was culturally classified as Western participants whereas the second group was classified as Asian participants. A simple questionnaire was prepared, which were presented to the participants:

1. Who are you: an Asian/Western?
2. Which webpage you visit daily and why?
3. What is culture to you?



4. How can culture aid in understanding human behaviour on the Internet?
5. What are the cultural issues, which affects your feelings on the Internet?
6. How do you feel for personalized or customizable webpage?
7. How would you rate (1=least important, 5=most important) these five user’s cultural characteristics: a) Intuition b) Privacy c) Networking d) Safety e) Self?

The result for Item 01-06 are summarized in the table 1 below. The Item 07 is shown graphically. In a graph below, an x-axis represents the two different cultures (C and D) whereas the y-axis represents the rating (1=least important, 5=most important).

Table 1. The summary table prepared below from the users for each item.

Users	Item 01	Item 02	Item 03	Item 04	Item 05	Item 06
C1	Western	Google, I search a lot.	Culture to me is not culture to you.	Change colour, language, and design based on cultural preferences	ad-words, pop-ups, malwares, virus	I like it does not matter.
C2	Western	Facebook, I want to be in-touch.	It is not an easy question to answer; well I think that perhaps culture to me varies with time, people, places, and technology.	I think some culture do not like pop-ups menus and some may like it so it should be based on what one culture likes or dis likes	bad design	Personalization is good idea.
C3	don't know	Wikipedia, can't say	I think this question should be re-written. I do not really know.	improve GUI (Graphical User Interface)	poor UI	Who cares?
C4	Western	MySpace, I feel comfortable using MySpace than Facebook.	Change	effective UI, interaction, touch screen enabled	lack of Uniformity of text, color and graphics	I do like to customize my webpage sometimes.
C5	Western	Twitter, I am interesting in Tweeting rather than Facebooking.	way of life	Should be safe and secure	depends on my mood for what I am looking	No, I do not prefer to personalize.
D1	Asian	Orkut, I have lot of friends in there.	Culture is about adapting to our environment based on other cultures	It is very difficult question to answer.	Unnecessary advertisement, pop-ups	Sometimes, I change the skin colour or theme.
D2	Asian	Facebook, Google, News-websites, and several other webpage. I like to read online materials.	Culture is mixture of different taste	Effective user interface	Numerous issues, This is why I used ad-block program	I prefer customizable options in a webpage.
D3	Asian	Facebook, I like to read what others have to say.	Culture is just a culture	relatively low download time	I do not know, there is nothing that affects my feelings.	It is great to see webpage in a personal way.
D4	Asian	There are several website and I do not remember I have to check my browser history.	maybe technological change	more culturally driven games	Maybe when browser stops running or Internet connect is gone	I do not know what to say.
D5	Asian	Orkut, I like to chat with friends online.	a taste of differences in a good way	can't say; maybe an intuitive driven webpage will be great	I think entire contents that I read on the Internet affects: on my personal feelings.	It is hard to say what fits right in a webpage but customizable is a good option.

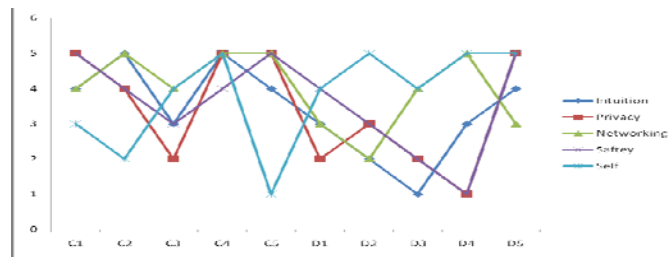


Figure 3. A graph demonstrating Item 07 from the research question.

### 4.1 Definition of Culture

During 1952, Alfred Kroeber and Clyde Kluckhohn gave 164 definitions on word: “culture” (Kroeber, A. L. and C. Kluckhohn 1952). In nonprofessional’s term the word ‘culture’ is defined as; “the way of life, especially the general customs and beliefs, of a particular group of people at a particular time”. Furthermore,

it has been suggested that there is no agreed definition of culture (HOFT 1996). However, as culture is very complicated in itself it would be extremely difficult to agree on common global definitions. In order to understand the definition of culture, we asked our participants about culture. We formulated following definitions based on their feedback.

Table 2. Definitions of Culture formulated based on user's feedback.

Users#	What users said (only in keywords)	Definitions formulated
1	<i>emotion, feelings, rules, ethics</i>	Culture is changing pattern of human values, emotions, feelings and thoughts with time based on person's location, ethics, rules and principles.
2	<i>behaviour, change, interest</i>	Culture is magnitude of adaptive behavior possessed by an individual, which changes dynamically based on self-interest.
3	<i>mixture of everything, social, ethical</i>	Culture is the mixture of various ingredients (personal, social, moral and ethical) unequal to national or international boundaries, race and religion.
4	<i>invent, fulfil needs, environment</i>	Culture is to invent, modify and update existing psychological and physiological needs based on external environment.
5	<i>ideas, learn from mistakes and pass it to others</i>	Culture is to learn novel ideas and theories, which are further pass on to younger generation. If 'n' is number of mistakes done in one's lifetime then 'n-1' is number of mistake which is passed on to younger generation.
6	<i>change</i>	Culture is change.
7	<i>create new things</i>	Culture is to create something new from nothing new.

## 4.2 Definition of Culture on the Internet

Nevertheless, the same general definition may not be true for users who interact on the Internet. The way someone behaves on the Internet with the applications and services could be dependent on many factors. Some of these factors could be the emotional state of the user, the purpose of browsing contents and profit users get for using the contents.

Table 3. Definitions of Culture on the Internet formulated based on user's feedback.

User#	What users said (only in keywords)	Definitions formulated
1	<i>daily habit, reading, writing</i>	Culture on the Internet is defined as a habit developed to facilitate online reading and writing skills.
2	<i>my needs, demands</i>	COI is defined as pursuit of user's demands by which services and applications should change on timely fashion.
3	<i>Facebook, Skype, Twitter</i>	COI is not only based on social networking experiences but also personal experiences of how a user uses different services and application on the Internet.
4	<i>mixture of everything, trust, privacy, security</i>	COI is defined as mixture of contents, software's, and applications, which not only demonstrate trust, privacy and security but also demonstrate pools of knowledge, communities and learning experiences.
5	<i>create, share, collaborate</i>	COI is about inventing novel tools and technologies on the Internet-based on experience and observations.
6	<i>culture inside cultures</i>	COI is defined as collection of different culture inside cultures.
7	<i>software, connected, computer networks</i>	Culture on the Internet is among users/software in the Internet, which is connected to computer networks within an individual user or group(s) of users/software.

## 5. CONCLUSIONS AND FUTURE WORK

This paper began by introducing the culture and culture on the Internet. The effects of Trompenaars, Hampden-Turner and Hofstede's cultural dimension were studied and considered. The findings demonstrate that two different cultures are very unlikely to change. We have developed a simple two cultural model in this study, which suggests that culture can vary based on the Environment. Two cultures could be same if and only if both of these cultures belong to the same environment. Moreover, two culture cannot be regarded same until one "adapts" to another culture based on given set of cultural characteristics. None of these cultures could be related if they do not consider user characteristics from one or multiple domain of cultures.

In this preliminary study, we have performed a pilot study with ten university participants from two different cultures and from two different universities with two different academic backgrounds. The participants addressed seven questions of which the last question was rated to test two-culture model. This simple model for testing two different cultures needs further study. Nevertheless, we believe that this is invariably significant contribution in understanding the cultures in HCI. We have learned that culture has greater impact in understanding how human behave on the Internet. Without the prior understanding of various models proposed by cultural scientists, it would be extremely difficult to conceptualize behavior of humans on the Internet. In addition, we also utilized Sanders and Dandavate model on the surface level, by Interviewing and collecting “explicit” knowledge from the users. Based on their feedback, we found that there is no hesitation for the division of culture such as Asian or Western. Meanwhile, they also suggested numerous definitions for culture and cultures on the Internet. The future work should include comparisons of human anxiety based on two different cultures from diverse cultural groups in detail.

## REFERENCES

- BARACK OBAMA, Barack Obama pleads for tolerance in Koran-burning row, 11 September 2010  
Available: <http://www.bbc.co.uk/news/world-us-canada-11265335>
- BASKERVILLE, R., Hofstede never studied culture. *Accounting, Organizations and Society*, 28, 1–14.
- HOFT, Ñ., 1996. In International User Interfaces. In: E. DEL GALDO AND J. NIELSON, Ed, New York: John Wiley & Sons.
- KROEBER, A. L. AND C. KLUCKHOHN, 1952. *Culture: A Critical Review of Concepts and Definitions*.
- MARCUS, A. & GOULD, E., 19 June 2000. Cultural dimensions and global web user-interface design: What? So what? Now what? *Sixth Conference on Human Factors and the Web*, 19 June 2000.
- RATNER, C. & HUI, L., 2003. Theoretical and Methodological Problems in Cross-cultural Psychology. *Journal of the Theory of Social Behavior*, 33, 67-94.
- TING-TOOMEY, S., 1988. Intercultural Conflicts Styles: A Face-negotiation Theory in Theories in Intercultural Communication. In: YOUNG YUN KIM AND WILLIAM B. GUDYKUNST, Ed, Newbury Park, CA: Sage, pp. 213-235.
- TROMPENAARS, F. & TURNER, C. H., 1997. *Riding the Waves of Culture: Understanding Cultural Diversity in Business*. New York: McGraw-Hill.
- Davis, F. (1993). User acceptance of information technology: system characteristics, user perception and Behavioral impacts. *International Journal of Man Machine Studies* 38, 475-487.
- Hubona, G. & Blanton, J. (1996). Evaluating system design features. *International Journal of Human Computer Studies* 44, 93-118.
- Hofstede, G. (1980) Culture’s consequences: International differences in work related values. Beverly Hills: Sage.
- Uren, E., Howard, R. & Perinotti, T. (1995). *Software Internationalization and Localization: An Introduction*. New York: Van Nostrand Reinhold
- Jeanette D. Snowball (2008), Measuring the Value of Culture: Methods and Examples in Cultural Economics, Springer
- Markussen, T & Krogh, PG (2008), 'Mapping Cultural Frame Shifting in Interaction Design with Blending Theory', *International Journal of Design*, vol 2, nr. 2, s. 5-17.

# 3D GESTURE RECOGNITION SYSTEM BASED ON DEPTH CAMERA

Yangkeun Ahn and Jiman Hong  
*Soongsil University*  
511 Sangdo Dong, Seoul, Republic of Korea

## ABSTRACT

Active research is underway on virtual touch screens that complement the physical limitations of conventional touch screens. This paper discusses a virtual touch screen that uses a multi-layer perceptron to recognize and control three-dimensional (3D) depth information from a time of flight (TOF) camera. This system extracts an object's area from the image input and compares it with the trajectory of the object, which is learned in advance, to recognize gestures. The system enables the maneuvering of content in virtual space by utilizing human actions.

## KEYWORDS

Gesture Recognition, Depth Sensor, Virtual Touch Screen

## 1. INTRODUCTION

This screens are being widely used at present owing to the ease of intuitive control. The touch screen, however, cannot be controlled if the production of a touch sensor is impossible (e.g. for large-sized screens), or if direct contact cannot be made in cases where the screen is located far away. To address these problems and facilitate control of touch screens, wide-ranging research has recently been carried out on virtual touch screens. Without any physical contact surface in place, a virtual touch screen generates a virtual screen at a given distance from the camera and recognizes the virtual touch of a physical object on the screen. Kim Hyung-joon and other researchers have proposed a dual camera-based virtual touch screen, which derives the three-dimensional (3D) position of a hand from images inputted from two fixed cameras and recognizes the touch point. Martin Tosas and Bai Li have implemented a virtual screen using a single webcam and a physical grid; when a hand is situated within the framework of the physical grid, as opposed to a virtual grid, a single webcam tracks the hand and recognizes a touch. Kim's research, however, implements a touch screen based on mathematical calculation of positioning, and as such it is subject to changes in the screen's position, as the screen is fixed. Tosas and Li fail to materialize a virtual screen, as 3D information is not included. Also, both studies implement limited touch features from among widely-varying human actions. A time of flight (TOF) technology-based camera provides 3D depth image information. The results can be expressed in black-and-white images; a virtual touch screen can be implemented in a simple and efficient manner by employing an image processing technique. A virtual touch screen, however, fails to take into account various human actions, as it simply realizes touch features based on touch points. Against this backdrop, this paper proposes a virtual screen that applies multi-layer perceptron technology to a virtual touch screen to recognize gestures. The structure of this paper is as follows: Chapter 2 explains a TOF camera-based virtual touch screen, and Chapter 3 discusses a virtual gesture screen based on the multi-layer perceptron. Test results are presented in Chapter 4; a conclusion and suggested directions for future research are given in Chapter 5.

## 2. TOF CAMERA-BASED VIRTUAL TOUCH SCREEN

A virtual screen can be easily implemented by utilizing the 3D depth information of a TOF camera and an image processing technique.

### 2.1 Depth Image

A TOF camera emits a laser or LED light. Using a built-in sensor, the time taken for the light molecules to return to the camera after touching the object are then recognized, and the distance from that object is calculated. The outputs of this camera are 3D depth images, which represent approximate 3D information.

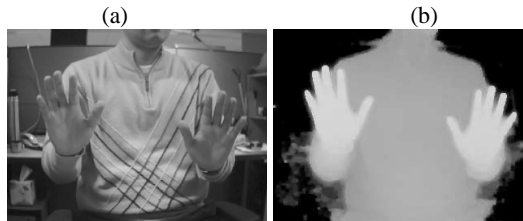


Figure 1. (a) Image input (b) Depth image

In Figure 1, (a) refers to the image input of the TOF camera and (b) the corresponding depth map. In (b), the depth map becomes whiter when the object is closer to the camera and blacker when it is farther away.

### 2.2 Setting of a Threshold Value

A virtual screen can be created by setting a given threshold value for the depth map.

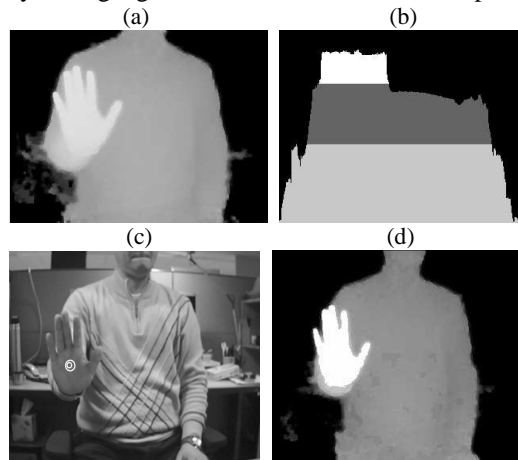


Figure 2. (a) Depth image (b) X-depth image (c) Object area tracking (d) Virtual screen image

In Figure 2, (a) is a 3D depth image and (b) an X-depth image, where the depth image is projected onto the X-axis. This X-depth image represents a bird's eye view on the camera. By setting a certain threshold value for the X-depth image, the area of a hand entering the virtual screen can be extracted, as shown in Fig. 2(d). Fig. 2(c) represents an image used in calculating the central moment of the hand area and tracking the hand area that reaches the virtual touch screen. This technique enables the implementation of a virtual screen.

## 3. VIRTUAL GESTURE SCREEN

A virtual gesture screen can be realized by incorporating an object trajectory database extraction system and a gesture recognition system into the implemented virtual touch screen.

### 3.1 Object Trajectory Database Extraction System

An object trajectory database extraction system consists of gesture input, gesture image correction, and feature extraction units.

#### 3.1.1 Gesture Input

Using a virtual touch screen, touched points are saved to create a gesture database. Whenever a touch action is performed, calculation is done at each frame to determine if the respective frames are touched. If touches are performed on a continued basis, the touched points are saved in the memory, and the gesture is assumed to have ended if an untouched frame is found. Then, as described in Figure 3, the points are connected to create an image.



Figure 3. Gesture input images

#### 3.1.2 Gesture Image Correction

When the gesture input is over, several phases should be carried out to correct the gesture image.

On this image, even the same gestures may take place in different positions and in differing sizes, and thus individual gestures need to be generalized. In other words, gesture recognition data regardless of position and size are required

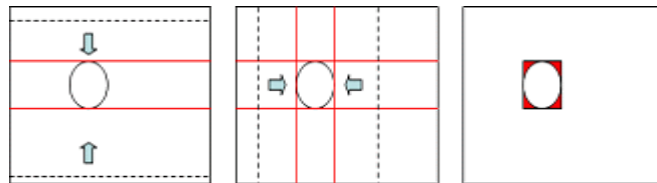


Figure 4. Positioning of endpoints at the edge of a gesture

For this purpose, the endpoints on the four sides of each gesture image are identified—as shown in Figure 4—and the image of the area, as described in Fig. 5, is projected onto a  $100 \times 100$  image. In other words, any image that is smaller than 100 pixels in width or length is enlarged, and one that has width or length greater than 100 pixels is reduced in size.

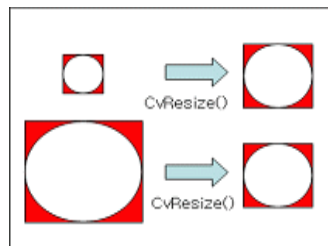


Figure 5. Readjustment of gesture image size

#### 3.1.3 Feature Extraction

After obtaining a gesture image with a size of  $100 \times 100$  pixels, the features to be used as the input unit in the perceptron should be extracted from the image. This paper divides an image of certain size into 25 smaller images with a size of  $20 \times 20$  pixels and, as illustrated in Fig. 6, the number of pixels in the gesture part of respective areas is taken as their features. Also, the target value of the perceptron is added to the end of the

database. If the total number of gestures is 3 and the inputted gesture is Gesture #1, the figure “1 0 0” is entered; if the inputted gesture is Gesture #2, “0 1 0” is added instead. This is because a sigmoid function is used during the learning process as an active function.

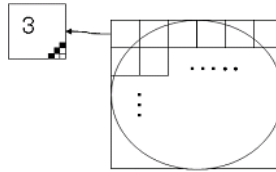


Figure 6. Number of pixels in the gesture part

## 3.2 Gesture Recognition System

The gesture recognition system is implemented based on multi-layer perceptron technology.

### 3.2.1 Multi-layer Perceptron

Multi-layer perceptron (MLP) is the representative algorithm for supervised learning. The algorithm brings the differences between output values obtained from inputs and predetermined targets of supervised learning (i.e. deviation) back to the perceptron structure and redistributes them in order to gradually reduce deviation levels during the learning process.

### 3.2.2 Structure of Perceptron

The multi-layer perceptron and data from the database explained in Chapter 2 are utilized to train the perceptron. In this case, 25 feature values, specified in 3.1.3, are used as inputs, and the learning target is set at the target value for the database; the number of hidden layers, learning rate, momentum, and target error rate are also determined in advance. Fig. 7 illustrates the structure of the multi-layer perceptron.

### 3.2.3 Gesture Distinction

After training the multi-layer perceptron, we worked on creating a gesture distinguisher. A real-time trajectory database program is run, and the data produced here are placed into the trained perceptron; its outputs are then used to distinguish gestures. The index of the node with the greatest nodal results for the output layer is used to distinguish a gesture from others. In other words, if the first node of the output layer is greater than the values of the other two nodes, the gesture may be concluded to be Gesture #1.

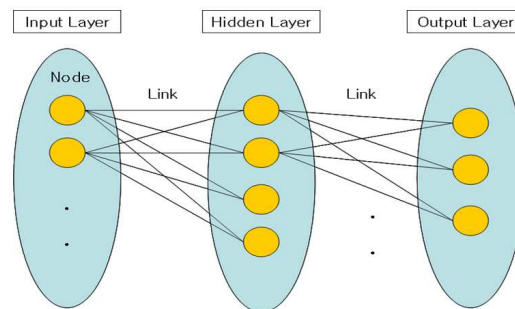


Figure 7. Structure of multi-layer perceptron

## 4. TEST RESULTS

To analyze the performance of this system, an experiment is performed to control a flash application with the three gestures described in Fig. 3. The number of hidden layers for the multi-layer perceptron is set at 10, the

learning rate at 0.1, and the momentum at 0.8. The target error rate is set to be 0.05 before initiating the learning process; five databases per gesture are used for the experiment. The experiment shows that the three gestures are different from each other and share no common aspects, and consequently less than 100 iterations of repetitive learning is sufficient to successfully distinguish the gestures. The flash program provides a menu selection feature when the user makes the first gesture of drawing a circle. The menu is turned to the right when the user draws a right arrow and to the left when the user makes the gesture of drawing a left arrow. Fig. 8 demonstrates how the flash application is controlled using the virtual gesture screen.

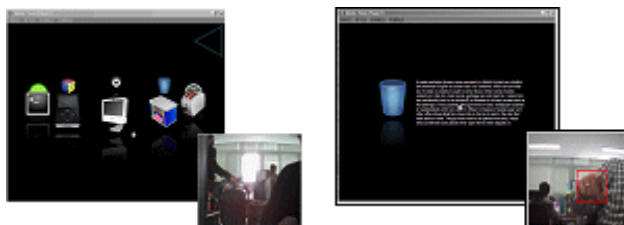


Figure 8. Controlling of flash program on virtual gesture screen

## 5. CONCLUSION

This paper has applied an object trajectory database extraction system and a multi-layer perceptron technology-based gesture recognizer to the conventional virtual screen system to enable gesture recognition as well as virtual touching. This makes it possible to express wide-ranging human actions, which can be expressed only in a limited manner through virtual touches in 3D space, and to develop a wide variety of content on this basis. The proposed system can recognize gestures on the virtual screen, but it identifies a single object on a real-time image and is applied only to a single gesture made by that object. The applicability of the system will be further enhanced if a multi-gesture recognition program is developed in the future that enables the recognition of gestures from multiple objects.

## REFERENCES

Journal

Kim Hyung-joon, 2006, Virtual Touch Screen System for Game Applications, *Journal of Korea Game Society*, Vol. 6, No. 3, pp. 77-86.

M. Fukumoto, Y. Suenaga, and K. Mase. 1994, Finger-pointer : Pointing interface by image processing. *Computers and Graphics*, Vol. 18, No. 5, 633-642,

Andrew Wu, Khurram Hassan-Shafique, 2002, Virtual three-dimensional blackboard: Three-dimensional Finger tracking with a single camera. *Applied Optics*, Vol. 43, No. 2, pp. 379-390.

Conference paper or contributed volume

Martin Tosas and Bai Li, 2004, Virtual Touch Screen for Mixed Reality ,*Lecture Notes in Computer Science*, Heidelberg, Springer Berlin, pp. 48-59.

Eunjin Koh, Jongho Won, and Changseok Bae, 2008, Vision-based Virtual Touch Screen Interface, *Proceeding of ICCE LasVegas*, USA,

Gokturk. S. B, Yalcin. H, and Bamji. C., 2004, A Time-Of-Flight Depth Sensor - System Description, Issues and Solutions, *CVPRW '04*, p35.

P. Hajela, B. Fu and L. Berke, 1992, Neural networks in structural analysis and design: an overview. *Computing Systems in Engineering*, pp. 525-538.

R.P. Lippmann, 1987, An introduction to computing with neural nets. *IEEE Acoust. Speech Signal Process 4 2* , pp. 4-22.

T. Darrell and A. Pentland. Space-time gestures. 1993, *In Proc. IEEE Computer Vision and Pattern Recognition (CVPR)*, New York, NY, pp. 335-340.

M. Hoch. 1998, A prototype system for intuitive Film planning. *In Automatic Face and Gesture Recognition*, Nara, Japan. pp. 504-509



# ENHANCING ZONE OF PRODUCT DESIGN WITH WIKI PRACTICE

Nikolay Yaremko

*Yandex*

*16, Leo Tolstoy St., Moscow 119021, Russia*

## ABSTRACT

This article examines similarities between wiki principles and modern design process. It introduces an idea of a “zone of product design” modeled after Vygotsky's zone of proximal development. This zone is understood as an area that team members explore by means of tools provided by designers, a certain level that is exclusively achieved as the result of combined effort. Designer's responsibility, in these terms, is to constantly keep the team in the zone of product design and support it as the concept of a product is being refined and deepened. It is demonstrated that group's capacity for reflexion and evaluation of its own activity is essential for working with the zone of product design. The main theses of the article are illustrated by concrete examples of wiki-like processes inside Yandex.

## KEYWORDS

Agile software development, agile practice, user experience design, product design, wiki, collaboration, cooperation, group consolidation, design methodology, zone of proximal development.

## 1. INTRODUCTION

The word *wiki* in the Hawaiian language means "fast". As Hawaiians tend to repeat words for emphasis, "very fast" would be *wiki-wiki*—which became the term of choice for Ward Cunningham as he was looking for a name for his programming patterns repository, the one that became the very first proper wiki around the Web (Cunningham, 2003). At present wiki is commonly associated with Wikipedia, a universally accessible and editable encyclopedia—allegedly able to meet offline competition both in terms of comprehension and trustworthiness (Lih, 2004); in fact, one of the most accredited source of knowledge around the Web.

For Cunningham himself (WikiDesignPrinciples, 2003), the word wiki primarily meant an approach to labor organization, a number of principles that he had devised—later to be practisized in most, if not all, existing wiki-systems. If we omit technicalities, the key wiki principles are as follows: *Simple*—easier to use than abuse; *Open*—should a page be found to be incomplete or poorly organized, any reader can edit it as they see fit; *Incremental*—pages can cite other pages, including pages that have not been written yet; *Organic*—the structure and text content of the site are open to editing and evolution; *Tolerant*— interpretable (even if undesirable) behavior is preferred to error messages and *Observable*—activity within the site can be watched and reviewed by any other visitor to the site.

Wiki principles are not confined to text-based environments; one could extend the definition to, say, free jazz music—a gradually built compound based on easy rhythm—sophisticated, organic and tolerant to deviations. This metaphor can be studied further in (Stambaugh, 2003). Wiki-gameplay could be defined as a multiplayer gameplay focused on repetitive and elaborate cooperative change of game rule set during the game itself (Yaremko, 2008). There is a couple of games embodying an idea of elaborate and cooperative construction of a gameplay in the process of a game itself: Jeep Form is a Live Action Roleplaying Game format built on collaborative improvisation (Wrigstad, 2008); various online forum games bring similar gameplay structure into the Web; Nomic, 1000 Blank White Cards, Week of Changes and some other games<sup>1</sup> are fine examples of board games implementing the concept of wiki-gameplay.

---

<sup>1</sup> [http://en.wikipedia.org/wiki/List\\_of\\_games\\_with\\_mutable\\_rules](http://en.wikipedia.org/wiki/List_of_games_with_mutable_rules)

One could also speculate that current trends in software development, namely shorter release cycles and agile techniques, are a good match for wiki principles. As Robert C. Martin notes, agile development is built upon three Cs: collaboration, communication and coordination (Martin, 2003). As design is often perceived a part of agile development, wiki principles, in theory, could be applied to design as well. This being said, it must also be noted that we recognize design as a process intended to form and model a concept of a product, specify the way it functions, implement the concept and finally introduce it to the market (Morris, 2009).

According to Kollman (Kollmann et al., 2008), in a typical agile project product design isn't something limited to the implementation phase, but rather a continual flow that keeps running through the life-cycle of a product. Every team member is included in this process—ranging from the management and marketing to the development engineers, who, by handling lots of back-burner issues, form an image of a product to a far greater degree than, for example, a consulting analyst who had laid out initial specifications. In such-like projects designers, as opposed to their accustomed role of feature-issuers, are challenged to support an open cooperative process, a wiki process, which includes the whole team as a source of shared product vision.

## 2. WIKIFYING PRODUCT DESIGN OF YANDEX.MAIL

For last two years wiki principles are being adopted in the development of the second largest webmail service in Russian market<sup>2</sup>. This practice leads not only to improved product quality and product-oriented motivation of the team, but also to reduction of implementation terms. In particular, recent major design change took about twice less time compared to previous three similar re-designs.

Found in 1997, Yandex is the leading internet company in Russia, operating the most popular search engine and most visited website ([www.yandex.ru](http://www.yandex.ru))<sup>3</sup>. In addition to search Yandex provides a broad range of internet services such as maps and navigation, shopping guide, etc. Yandex.Mail is another key product of Yandex, webmail service is second only to Mail.Ru. The company has over 3000 employees worldwide to date and Yandex.Mail product team consists of 45 people including 4 project managers, 6 server-side interface developers, 3 client-side developers and 3 user experience specialists.

Yandex.Mail interface is being routinely optimized and augmented with new features for almost eleven years of its existence. Major redesigns accompanied by technology and user experience changes were released approximately each 3 years: in 2003, 2006 and 2009. Each release was designed and implemented in about 1.5–2 years separately from local improvement routine. It was easy to measure effect of local improvements with controlled experiments similar to described by Kohavy et al. (Kohavy et al., 2008), while major experience change was monitored mostly by *decline rate*: a percentage of users switching back to previous interface after they give a try to a new one. Successful enough redesigns score 4–5% declines of daily active audience, this leads to less than one per cent decline among whole audience.

Decline rate is not a good meter for experimenting with design decisions thus product team members usually rely on their own vision of the product. Moreover, almost everyone in company is familiar with webmail and have his or her own opinion on the product's interface, look and feel and even strategy. Openness and democracy declared in Yandex's code of conduct<sup>4</sup> require product team to listen up to these opinions regardless their author's status thus product team is required to take in consideration such visions not only of top managers but of many other colleagues.

Next major technological upgrade was started in 2009. Product team was aiming to dramatically increase interaction speed and responsivity of interface by means of both AJAX technology and experience design (this goal was achieved at whole). It was also required to incorporate several companion products such as Calendar and RSS Reader in a solid user experience of Yandex.Mail. Product team size for this project was the very same as for the previous major release: two managers, seven interface developers (including two on client-side), a contractor visual designer and in-house user experience designer. This situation seemed to be a fruitful soil for wiki principle implementation thus several practices were introduced. Designers' role in most of these practice was significant if not crucial.

<sup>2</sup> Yandex.Mail serves 4.5 million daily and 14 million monthly visitors while Mail.Ru reaches 11 and 23 million daily and monthly visitors accordingly. These numbers are provided by TNS Web Index and published by both competitors online.

<sup>3</sup> In 2010 Yandex generated 64% of all search traffic in Russia, its homepage attracted a monthly average of 21.5 million users, and Yandex were the largest internet company in Russia by revenue: <http://about.yandex.com/>.

<sup>4</sup> Some excerpt of the Yandex code of conduct in Russian: <http://company.yandex.ru/about/main/>.

**Incremental and organic process** of development and deployment was organized by series of beta releases with trimmed functionality: first releases were inside-only, after a while limited and then public beta-testing was open. The incremental release practice was introduced by product manager and the function of a designer was to help product team with identification of key product functions and scenarios. First beta was made available for 3000 people of Yandex staff right after three key parts of interface were ready: messages list, message view and compose form. Other parts were under development and some of them such as settings pane and address book were not even sketched by designer. Invites for limited beta-testing were sent outside three months earlier than address book was implemented in a new design.

Designers' work was organized accordingly; for instance, cross-system detailed interface specifications were replaced by interactive prototypes and short guidelines. Since previous re-design the team was provided with a detailed specification of how interface response notifications should work. However—mainly because of the length of the document—bringing it to everyday use was anything but simple. Naturally, developers preferred to halt every time they had to deal with the notification system and ask for advice. Based on discussions with the team, the designer drafted a simple guideline, together with a number of prototypes that illustrated its major points. After awhile, by joint efforts of the whole team, a new specification was born, essentially rendering the prototypes useless for the developers. Curiously enough, this new notification system proved to be significantly less intrusive as compared to the original specifications.

**Open process of communication about product issues** was organized in the form of regular (1–2 times a week) meeting on design implementation. Meeting agenda was gathered collaboratively using corporate intranet which is also wiki-based—everyone in the company was able to add his or her question, suggestion or a problem to that agenda. The meeting itself was open for product team—those who would like to know how and why product decisions are made were welcomed at the meeting; only presence of product manager and designers was mandatory. This meeting was used to gather and discuss any product ideas from the team. By means of this process developers were gaining knowledge of the logic hidden behind the curtain of solutions provided by designer; their own post-hoc opinion was that such a process was crucial for their motivation. As the meeting was scheduled and driven by the needs of developers themselves, designer was expected to discover incentives which prompted his solutions thus developers were able to reproduce missing details on their own. On other hand designer's function was to assimilate ideas proposed by other team members and to visualize such ideas providing whole team with deeper perception of each other vision.

**Observable product design** was discussed by means of corporate social network provided a place for open communication and displayed behavior. A discussion club dedicated to the mail service includes 223 persons which is approximately 15% of the total number of people involved in product development at whole company (five times more than the size of product team). Only about a half of them were mere bystanders who have never participated in discussions. For the last two years, this club had 1674 posts in 302 discussions with the average of 7.85 comments in each. The majority of graphical layouts, concept notes, prototypes and early betas became immediately available for discussion in this club.

Designers brought life to that discussion club: they picked out mockups and prototypes for publication in order to reveal current directions of product growth to the team while avoiding portrayal of chores. Product team lead by designers collected peer reviews to find fresh ideas and avoid exhausting approvals. By these means Yandex, being a large company, provided an environment where by the time of product presentation all the parties involved in a project have a notion of its essence, key features and interface.

**Tolerance** was supported in the process of opinion gathering: while designer acted as a moderator, he welcomed substantive outgivings and stood back from it appraisal. Designer initiated dialogues to study out particular user experience issues (e.g. interface behavior or apparel details) that led to overall impression. This behavior displayed in public allowed designers to spread culture of mutual tolerance: product team members along with other spectators moved on from a position of pundits and censors and became active contributors. Instead of keeping tabs on 'designer errors' fellow developers began to describe technical subtleties that obstruct implementation.

**Perception simplicity** was a governing principle in designer routine: any mockup shown to the product team should be as close to a final product in terms of look and feel as possible. About 42% of layouts featured in the Yandex Mail discussion club were accompanied by HTML prototypes. These prototypes allowed to perceive a mockup as a ready-made product, completely eliminating the constraining attitude frequent in these cases. As a result, comments grew more substantive and came in greater numbers.

**Finally**, present version of Yandex.Mail was delivered seven months after the first interface mockup was accepted by the product team; the very same process took more than a year for previous major release. We

relate it to improved communication amongst designers and product team. Most benefit was reaped on early stages of development because of lesser volume of solution specifications and more effective process of iterative follow-up revision: total development time became 2.2 times shorter compared to previous version.

Decline rate was below 1%— much lower than 4–5% decline rate of previous major release. Given the product team was merely unchanged we propose that the introduction of wiki practice made positive influence on both the product quality and the time of development and adoption. Subjective feedback gathered from product team members allowed us to consider growth of their motivation; this speculation correlated with a raised number of product enhancements proposed by the team.

### 3. ZONE OF PRODUCT DESIGN

It is our understanding that designers—along with their apparatus, such as user scenarios, specifications, guidelines, mockups and experiments—constantly form a fluid “zone of product design”, i.e. an assortment of issues that the team can't resolve by itself, but easily handles when offered designer's help. This name has been chosen as a paraphrase for the term “zone of proximal development” introduced in (Vygotsky, 1934), published by Harvard University Press in 1978. Vygotsky defines this zone as “an area of maturing processes that includes tasks a child is able to accomplish only if supported by an adult; a certain level that is exclusively achieved as the result of combined effort”. Much like children leaving this zone as their activity becomes more and more complex, product development team would come out of the zone of product design and stop seeking designer's help when it's no longer required. Instantly a new scope of tasks is shaped, and the helping hand is called for. Designer's responsibility, in these terms, is to constantly keep the team in the zone of product design and support it as the concept of a product is being refined and deepened. This perception of designer's function is, of course, subject to debate.

Let us consider specifics of designers' relationships with the zone of product design. First of all, they shape this zone, i.e. look for questions, issues and tasks that the team can't resolve by itself, but will be able to with designer's help provided in form of specifications, mockups and blueprints. Secondly, in a manner of speaking, designers constantly sharpen their tools, i.e. create aforementioned specifications, wireframes, site maps, flow charts and other outlines of various abstraction levels; develop mockups and interactive prototypes that demonstrate how the project or its most important parts should work; prepare and carry out experiments; do split-tests and collect statistics that the team can use to verify different hypotheses.

Another important task is to expand the available zone of product design. At any given time every team member has an area that they understand by means of tools provided by designers. This area is a Z.P.D. The broader this area is, the easier it is to fine-tune one's understanding—with designer's help. The width of the zone is determined by how quickly and readily each member of a team accepts the game—that is, tries to consider and assimilate new points of view. Works of Vygotsky (Vygotsky, 1962; Vygotsky, 1978) provide some insight into psychology of interaction and group work.

There's little or no need to maintain this process, however, because, as pointed out by Tinsley & Lebak (Tinsley & Lebak, 2009), any group of people working on the same project for a long time, would by itself shape a so called Zone of Reflective Capacity—a special case of Vygotsky's Z.P.D.s. Inevitably, any such group will increase its capacity for reflexion and evaluation of its own activity. So as trivial as it sounds, for a group to be more creative and effective it must be consolidated—something that a designer might want to take notice of on a day-to-day basis. Being a creative professional, a designer is able to offer a knee to other members by transforming their judgement into graphical and interactive artefacts, i.e. mockups and prototypes. Shifting focus from people to mockups helps separate personal attitudes from decision-making; thus allowing resolutions to be made promptly and thoughtfully.

### 4. DISCUSSION AND FURTHER WORK

According to our reckoning, designers should not usurp conceptual work on product design. Instead, they should organize a collaborative process of generating collective thought. In order to avoid ‘design by committee’ effect decisions should be made by product manager and product designer while the product team should be given a toolbox that will allow them to understand the project on a deeper and more precise

plane. In addition of collective thought visualization, the function of a designer is to be a wiki-culture extensionist—designer's behavior could become an easy act to follow due to demonstrative character of his or her work. Wiki design practice implies that the graphics is produced as openly and progressively as possible; the result of these efforts is instantly presented to the team. The team, in its turn, vows to tolerate mistakes and inaccuracy while trying to improve designer's work by means of discussion and constructive criticism.

We have observed that collective thought on product is poorly compatible with routine of task designation and execution. 'Manager' and 'executor' roles are required for such a routine, while any institutionalized order of command creates inequality thence obstructs group consolidation. Numerous conflicts we observed between managers and developers seem to attest this proposition. Discussion on possible improvement of the product is usually turbulent and also leads to a breakdown or even a complete stop of routine task execution. Voluntarily used corporate social network allows to separate an open discussion club from traditional workflow and to avoid possible disruption of routine operations. Both online and offline discussions were announced as rank-free and regalia-ignorant in a way that equal treatment was guaranteed for each disputant.

The Yandex.Mail case of wikifying design processes was recognized as a success by company staff, therefore wiki practice is being introduced to a number of projects inside Yandex. For example, new music streaming service development was supplied with online discussion club with 209 participants and 217 active discussions. At the state-of-art it appears that practice blossoms if both project lead and designer are involved. We intend to investigate implementation of wiki practice further in order to convert specific tricks into techniques and then to build up a solid methodology of how wiki design process should be organized.

It is our perception that product is crafted while team is working in the zone of product design formed by means of designers' work. In order to help designers to extend their responsibility for this zone we intend to refine a set of specific organisational tools and techniques based on organisational activity games (Shchedrovitskii & Kotel'nikov, 1988). The methodology of such a research is also a subject to further discussion for it is not quite clear how to measure outcome of a complex organisational practice.

## REFERENCES

- Beck, K. and Ralph, J., 1994. Patterns Generates Architectures. *Proceedings of European Conference of Object-Oriented Programming*. Bologna, Italy, pp. 139-149.
- Cunningham, W., 2003. *Correspondence on the Etymology of Wiki*. WikiWikiWeb, <http://c2.com/doc/etymology.html>.
- Cunningham, W., 2003. *Wiki Design Principles*. Portland Pattern Repository, <http://c2.com/cgi/wiki?WikiDesignPrinciples>.
- Kohavi, R. et al., 2008. Controlled experiments on the web: survey and practical guide. *Data Mining And Knowledge Discovery*, Vol. 18, No. 1, pp 140–181. Springer Netherlands, Netherlands.
- Kollmann, J. et al., 2009. The Importance of Identity and Vision to User Experience Designers on Agile Projects. *2009 Agile Conference Proceedings*. Chicago, USA. pp.11-18.
- Lih, A., 2004. Wikipedia as participatory journalism: Reliable sources? Metrics for evaluating collaborative media as a news resource. *Proceedings of the International Symposium on Online Journalism*. Austin, USA.
- Martin, R. C., 2003. *Agile software development: principles, patterns, and practices*. Prentice Hall, Upper Saddle River, USA.
- Shchedrovitskii, G. P., Kotel'nikov, S.I., 1988. Organisational activity games – a new way of organising and a method for developing collective thinking activity. *Soviet Psychology*, Vol. 26.
- Stambaugh, T., 2003. *Jazz Music Metaphor*. Portland Pattern Repository, <http://c2.com/cgi/wiki?JazzMusicMetaphor>.
- Tinsley, R., Lebak, K., 2009. Expanding the Zone of Reflective Capacity: Taking Separate Journeys Together. *Networks* Vol. 11, No. 2.
- Vygotsky, L. S., 1934. *Thinking and Speech*. Published in 1962, The M.I.T. Press, Cambridge, USA.
- Vygotsky, L. S., 1978. *Mind and society: The development of higher psychological processes*. Harvard University Press, Cambridge, USA.
- Wrigstad, T., 2008. The Nuts and Bolts of Jeepform. *Playground Worlds Creating and Evaluating Experiences of Role-Playing Games*. Ropecon, Solmukohta, Finland.
- Yaremko, N., 2008. *Define Wiki Gaming*. Retrieved from <http://wikisym2008.humanemagica.com/wikigame.html>.

# TOUCH-BASED INTERACTION: A PROPOSAL FOR AAL

J. Bravo, R. Hervás and J. Fontecha  
*MAM I Research Lab at Castilla-La Mancha University (Spain)*  
*Paseo de la Universidad, 4 - 13071-C. Real (Spain)*

## ABSTRACT

There is no doubt that cell phones are the most widespread electronic devices all over the world. However, they are underused and no more than 15 % of their potential serves the user. In this work we analyse the possibilities to help elderly people by using this device and trying to minimise the needed interaction with it. To this end, we propose the use of Near Field Communication for supporting daily life problems of elderly people living alone at home.

## KEYWORDS

Touch Based Interaction, NFC, AAL.

## 1. INTRODUCTION

Nowadays the population is aging at an alarming rate worldwide. As a consequence, some problems like independence, safety, difficult relationship with technologies for supporting the daily routine by cognitive declining are problems that governments are trying to solve with many initiatives because the progressive aging of people has economical consequences, not only for future decades, but currently.

Ambient Assisted Living (AAL) promotes the use of technologies for helping elderly people in order to maintain their autonomy, increasing quality of life and facilitating their daily activities for augmenting the time living alone at their homes. For this reason, we have considered mobility as one of the most important problems for AAL [1]. In this line, the European Commission encourages technology and innovation for AAL. Thus, the research community in this area believes it is important to adapt technologies for supporting elderly people living alone at home. However, these technologies must be closer to these kinds of people, minimising the user interaction, even trying to eliminate it. The reason, as we mentioned before, is elderly are not familiar with technologies. For that reason it would be desirable an interaction as simple as possible.

Bearing in mind all described above, in this work we propose a solution to improve the care-dependent people's autonomy by using cell phones, the most widespread devices all over the world, and (NFC), a short-range communication technology. NFC-based cell phones can communicate with small tags attached to daily objects. The communication automatically starts when the cell phone gets close to the tag. This way, the interaction with the environment is fast, natural and intuitive because the user receives the services by just touching an object with his cell phone. This fact represents a solution to help care-dependent people for carrying out their daily activities. Moreover, since the user interacts with the environment in an easy and natural way, elderly people's reticence to use new technologies will not be a problem.

There are some approaches by using technologies for supporting daily activities. In [2] the authors propose an NFC-based system for supporting elderly people to choose their meals to be delivered by means of a home care service. In [3] a proposal for NFC-based assisted living with a shopping list assistant is presented. Also, the shopping list could be sent to relatives by touching a photo. Although this approach is similar to our proposal, it is only for the shopping list service and does not contemplate the idea of the scalability for offering more services to elderly people. Our experience for supporting the elderly by adapting NFC technology is varied. In [4] and [5] the touch-based interaction concept is presented. In the next section we present the NFC and touch-based interaction. In Section three NFC-based solutions for shopping list, catering service and mobile prescription are presented. In Section four a generalisation of these kinds of services called services catalogues is proposed. Finally, we draw some conclusions and future work.

## 2. AIMS AND SCOPE: TOUCH-BASED INTERACTION

Simple interaction is a key concept of our system that should be made. Elderly are not familiar with new technologies, so it would be desirable to have an interaction as simple as possible [5]. In order to create applications for supporting daily activities in an assisted environment, it is necessary to adapt sensorial capabilities. From our experience, keeping in mind Weiser's vision, we have applied identification technologies. Thus, we can achieve this objective using a Bluetooth enabled NFC cell phone. The interaction is as simple as bringing the cell phone close to the device, object or person with which they want to make any action. For example, vital signs through monitoring devices, displays, computers, photos or pictures, etc. can be managed easily.

For all the above-mentioned, we propose the adaptability of the NFC. It is a short-range technology using a frequency of 13.56Mhz and consists of two elements: initiator and target. The first one controls the information exchange and the second one responds to the initiator. In this technology there are available two operation modes: active and passive. In the first case, both devices generate their own field of radio frequency transmitting data and, in the second one, there are modulating data transfers. Finally, it is important to mention the four modes of operation depend on the two devices implicated transferring information are available: cell phone-cell phone, cell phone-reader, tag-reader and cell phone-tag. In our proposal, we only use the last one, it means, cell phone and tag. Note that tags can be attached to daily objects.

Our system must be adapted according to the abilities of care-dependent people. Thus, user interaction has to be as simple and intuitive as possible. To that end, we have adapted NFC technology. The user interacts with the system by touching NFC tags, which are scattered all over the environment, with their cell phone. Each tag stores specific contextual information that enables the phone to offer the required services to the user. We are talking about a touch-based interaction.

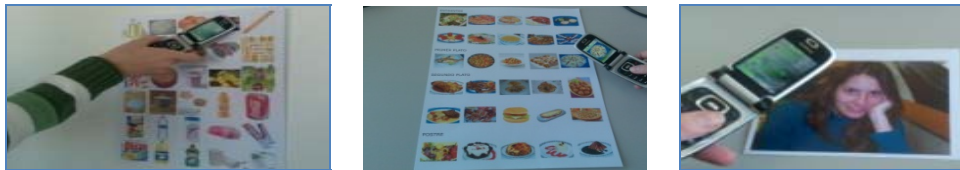


Figure 1. Applications for shopping list, catering service and relatives' calls

We have developed four applications. The first one refers to the shopping list. Some people find barriers going to the supermarket and need assistance for carrying out the daily needs of food or articles for cleaning the home. In this sense we have prepared carefully a product catalogue with a number of tags covered by the corresponding icons as can see on the left of Figure 1. Once the tag is touched, the assistant receives in the cell phone the new petitions to buy at the supermarket. The second picture shows the catering service, including a similar catalogue containing dishes. Users, choosing their preferred menus, receive their orders at home no more than 30 minutes later. The third one refers to phone calls to relatives and friends. For that, we place tags in the photo gallery and automatically, touching them, the dialling is opened. Finally, the last application (not in the figure) corresponds to the medical prescriptions. This process is the most important for us because users must visit the doctor to get prescriptions and, as we have mentioned before, mobility is a problem for elderly people. This application is detailed in the next point.

### 2.1 Mobile Prescription

In order to obtain the medical prescriptions, elderly people have to phone the healthcare centre to arrange an appointment with the doctor. After that, they have to go there, wait for their turn to come and go back home with the prescription papers. All these activities make care-dependent people's life even more difficult, especially of those who have mobility problems. Mobile prescription claims to improve care-dependent people's quality of life. To that end, patients will be able to obtain the prescriptions of their medicines from home, without having to go to the healthcare centre.

As we said before, to achieve an easy user interaction, we have used NFC technology. Each medicine box contains an NFC tag that stores information relative to the drug. The patient uses an NFC-enabled cell phone

and has an NFC tags panel to interact with the application installed on the phone. Each NFC tag of the panel is stuck under an intuitive picture that represents its function, that is, the service that the cell phone offers when it detects that tag. In this way, for example, when the user wants to see their prescriptions, they just have to touch the picture “See my prescriptions” with their cell phone; consequently, the application is automatically launched and it shows the patient’s prescriptions. This simple interaction spares the user from manually starting the application and selecting the option “See my prescriptions” in an options menu.

Many care-dependent people might have eye diseases, so we have adapted the system to these people’s needs. The application installed in the user’s phone not only informs the patient showing information on the screen, but also plays a voice that says the information that is shown on it. To that end, we listed the different messages the application has to show. Using a text-to-speech application, we created the different files the phone might play. Each time the application has to inform the user, it shows the message and plays the corresponding sound file.

## 2.2 Tag Catalogues

Once we have built the above-mentioned applications, our proposal is basically to generalise these services and others that can be implemented. For it, we propose the architecture shown in Figure 2. As we have commented before, people at home only have to touch the corresponding tags with their cell phone in order to obtain services. For that, we have developed a cell phone middleware responsible for the 3G transmissions. In it, on the one hand, some data such as server IP or phone number are included, each with the user ID. On the other hand, data about the service is required, that is, the service code, product units and server replay (check). Therefore, the functionality of the middleware is reduced to the data transmission.

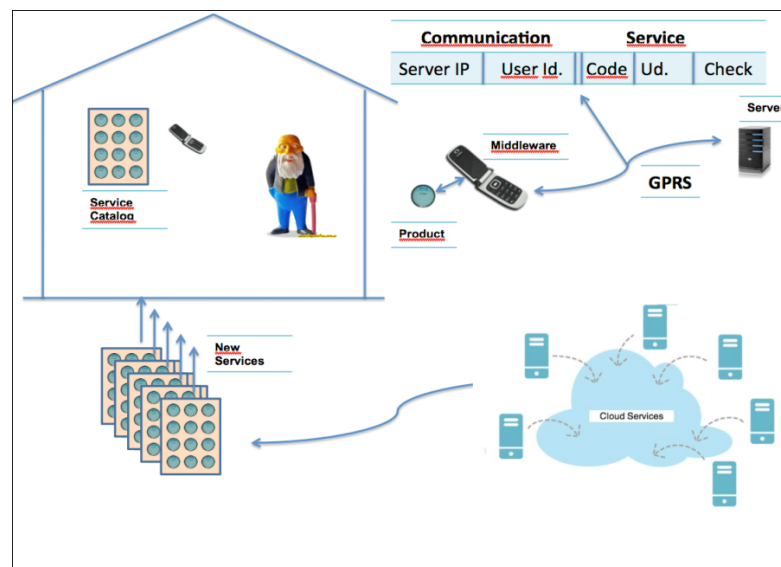


Figure 2. Architecture for AAL daily needs.

In order to achieve a scalable system, this middleware has to support new services, for that, it is necessary to standardise the communication protocol. Doing that, all possible solutions can be included in a cloud services. At this moment, providers can download the required server applications. With them, they are allowed to manage their customer transmissions, printing and personalising their tags’ catalogues. For example a home repairs service can be implemented easily.

Table 1 shows an example of transmissions with the three mentioned solutions. In it we can see takeaway (TA), shopping list (SL) and pharmacy (PH) examples. Additionally, product descriptions with the identification and check control can be observed. Note that in the case of pharmacy, the check control is only received by GPRS from the doctor server as a combination of user and product identification. This check will be necessary for obtaining medicines at a pharmacy.



For completing this proposal, a solution for providers is needed. In this sense, we pretend to support the process for making tags' catalogue services. In it, it is important to consider the way in which the transmission process is realised because it has to be adapted to the mentioned middleware. Also, the idea of generalised services is another important concept.

Table 1. Data Communication Structure

Server IP	User Id.	Product Id.	Ud.	Check
167.165.76.32	1221450012	TA-Chicken-17001	1	Received? ok
167.165.76.32	1221450012	TA-Salad-17005	1	Received? ok
167.165.76.32	1221450012	TA-Banana-17011	1	Received? ok
192.181.92.11	1221450012	SL-Bread-29087	1	Received? NO
189.178.11.43	1221450012	PH-Gelocatil-87576	1	875761221450012

### 3. CONCLUSION

We have studied the elderly people barriers for living alone by using simple technologies like NFC. First of all we have implemented some applications reducing the interaction by only touching tags. Then, a generalisation is proposed through the combination of middleware, communication protocol and server application.

### ACKNOWLEDGEMENT

This work has been financed by PIII09-0123-2762 project of JCCM (Spain).

### REFERENCES

- [1] The Ambient Assisted Living (AAL) Joint Programme, <http://www.aal-europe.eu/>
- [2] Häikiö, J. et al (2007). "Touch-Based User Interface for Elderly Users". Mobile HCI'07, Singapore.
- [3] Llorente, M.A. et al (2009). "Analysis, Development and Evaluation of AAL Applications Based on NFC Technology to Support Independent Daily Life". M. Tscheligi et al (Eds.). Aml 2009. Salzburg (Austria).
- [4] Bravo, J., Hervás, R., Fuentes, C., Villarreal, V., Chavira, G., Nava, S., Fontecha, J., Casero, G., Peña, R. and Vergara M. (2009). "From Implicit to Touching Interaction by Identification Technologies: Towards Tagging Context". 13th International Conference on Human-Computer Interaction HCII 2009. San Diego, CA, USA. July, 2009.
- [5] Bravo, J., Hervás, R., Chavira, G., Nava, S. and Villarreal, V. (2008). "From Implicit to Touching Interaction: RFID and NFC Approaches". Human System Interaction Conference (HSI'08). Krakow (Poland).
- [6] Bravo, J., López-de-Ipiña, D., Fuentes, C., Hervás, R., Peña, R., Vergara, M., and Casero G. (2008). "Enabling NFC Technology for Supporting Chronic Diseases: A Proposal for Alzheimer Caregivers". In European Conference, Aml 2008. LNCS, Springer. ISBN: 978-3-540-89616-6. pp. 109–125. Nuremberg, Germany.

# INTERACTIVE SONIFICATION OF MEDICAL DATA

Stefan Kovar and Margit Pohl

*Vienna University of Technology, Institute for Design and Assessment of Technology  
Favoritenstr. 9-11, Vienna, Austria*

## ABSTRACT

In this paper we describe the development of a sonification system to represent time-dependent data. The system's aim is to support therapists of anorexic young women to analyze questionnaires filled in by their patients. The mapping we adopted was suggested by the therapists and reflects the "mood" of the variables we used in the sonification (e.g. depression). Our main questions were, whether sonification is applicable in this context, whether the mapping we chose was appropriate and whether users get an overall impression of the patients' development. The results of this pilot study indicates that our approach is promising.

## KEYWORDS

Sonification, time-dependent data, evaluation, information interfaces.

## 1. INTRODUCTION

In recent years, the ability of computers to process and present data has increased tremendously. Information visualization is applied to give users a concise overview of data, but they are often overwhelmed by the sheer amount of information on the screen. Sonification might be used to alleviate this problem and complement information visualization. It should be pointed out, however, that there is still not very much information about the potential benefits of this combination.

In this paper, we describe the development and evaluation of a sonification tool to support therapists analyzing time-dependent data about their anorexic patients. The data used for this study was previously used for the development of an information visualization tool (Rester et al 2007). In the course of the previous study, we realized that the potential benefit of the visualization system might be enhanced by sonification giving the users a more comprehensive overview of the data. We decided to first investigate whether sonification might, in principle, be an advantageous solution to the problem of information overload in our context. Therefore, we developed a stand alone system without combining it with an information visualization tool.

The system we developed was aimed at supporting therapists in their work with anorexic young women. These patients, and also their relatives, have to fill in many questionnaires before, during and after the therapy. Interesting variables in this context might be, e.g., depression or number of friends the patients have. The therapists need these data to clarify which factors influence the success or failure of the therapy and to predict the outcome of the therapy process. Time-dependent data play an important role in this context. In sonification, the change of variables over time can be displayed quite naturally (Flowers 2005).

In this paper, we first discuss related work. Then we describe the system we developed. This system presents information from four different questionnaires (that is, four different sounds) at the same time. In this way, users not only get information about single values, but should also get an overall impression of the patient's development. We also tried to develop a mapping which is meaningful to the therapists working with the patients (e.g. dissonance as indicator for a negative development). In the last section, we describe a short usability study to assess whether this approach works and potential users can derive insights from such a tool. The main contribution from this work is to investigate a sonification system presenting variables in parallel streams. This will be of interest to designers of sonification systems as well as to persons wanting to analyze complex questionnaire data.

## 2. RELATED WORK

Traditionally, sonification is used to support visually impaired people. In recent years, the combination of visual and auditory stimuli has been investigated, either to ease the problem of information overload (see e.g. Salvador et al 2005) or to study how design principles from information visualization might be applied to sonification (see e.g. Zhao et al 2008). Salvador et al (2005) show that the combination of sound and InfoVis methodologies can yield better results than both methodologies used individually.

Brewster (2002) lists several advantages this combination of output modalities in computer interfaces could have. Among others, he mentions that auditory information display can reduce the cognitive load on the visual channel. Walker and Mauney (2010), e.g., point out that one of the most important issues in sonification is the mapping of variables to sound dimensions. Some guidelines exist which are valuable for the design of sonification systems (see e.g. Flowers 2005, Brown et al 2003). Flowers (2005) gives a comprehensive overview about design guidelines for sonification. He especially outlines which features of a sound may be used for representing what kind of data. In this context, he points out that the sensitivity of the auditory system for temporal changes is very high, and that such changes sometimes might be better represented in an auditory form than in a visual form. He also suggests not using loudness to represent continuous variables. Brown et al (2003) also developed a fairly comprehensive system of guidelines. In their studies they found out that users prefer parallel mode to sequential mode of data presentation and that this mode was also more efficient.

The guidelines developed by Flowers (2005) and Brown et al (2003) informed the design of our system, especially the recommendation to use parallel mode, but to design the system carefully to enable users to distinguish the different variables. Flowers (2005) points out that there are still no obvious solutions in this area and more research is necessary. Recently, some sonification systems were developed to support scientists in the analysis of large amounts of data (see e.g. de Campo et al 2006). These efforts seem to be quite successful and interesting and inspired us in our work. A longterm goal of our work is also to investigate the combination of visualization and sonification, although it is not a specific issue in this paper.

## 3. DESCRIPTION OF THE INVESTIGATION

### 3.1 The Sonification System

The sonification system was developed in close cooperation with the therapists from the hospital. They also provided the data. The system represents data from questionnaires which were presented to therapy patients five times (before, three times during, and after the therapy). The therapy lasted approximately one year. For the purpose of the development of the sonification system we used the data of six patients and four representative questionnaires:

1. Beck-Depression Inventory (depression)
2. Phobia Scale (anxiety in social situations)
3. Self-Efficacy (whether patients believe they can act effectively and reach their goals)
4. Euthymia (whether patients enjoy rewarding and relaxing activities)

The first prototype approach was to make the mapping of questionnaire values and sounds as simple as possible. The different questionnaires were represented by various timbres, the single values by pitch. We developed a small prototype which we showed to the therapists to get feedback from them about the mapping of the variables. The therapists suggested using different metaphors for the sonification to correspond with their perception of the patients. They argued that a more meaningful mapping could support them better in the process of data interpretation. They both came to very similar suggestions. Depressive patients, e.g., should be described by musical sounds of low volume in a minor key. It should be pointed out, however, that sounds in a low volume, if presented in combination with other sounds may not be audible. This has to be taken into account when designing such an interface. In general, the therapists proposed that patients with a positive development should be represented by harmonious sounds and patients with a negative development by dissonant sounds. We discarded the first prototype and developed the final system based on the therapists' suggestions. The following mapping was developed:

- Beck-Depression Inventory (measuring depression)
  - positive: major key, loud
  - negative: minor key, low
- Social Phobia Scale
  - positive: low-pitched, loud
  - negative: low-pitch, low/hushed
- Self-Efficacy
  - positive: fast tempo (rhythm)
  - negative: slow tempo (rhythm)
- Euthymia (whether patients enjoy rewarding and relaxing activities)
  - positive: light timbre, gay
  - negative: dark timbre, dissonant, bleak

Each dimension consists of a scale ranging from 1 to 5. The mapping reflects this structure of the scales and presents different sounds for every point of the scale. The sounds representing the four dimensions are presented in parallel. This is recommended by e.g. Brown et al (2003). The development in time is presented in a discrete fashion. The users may interact with the system to decide when and how long they want to hear a sound.

The sonification system was developed using the software SuperCollider for the Mac. For the evaluation (see next section), a graphical interface was developed, also using SuperCollider (see Figure 1). The interface consists of different buttons. There is one set of buttons in the top right corner presenting single sounds when pressed. These are the reference sounds of the four different scales (e.g. SPS is the Social Phobia Scale). All other buttons present composite sounds, that is sounds consisting of four individual sounds. In the upper half of the screen are test sounds for the learning phase, in the lower half are sounds representing the tasks of the test phase.

### 3.2 Evaluation

In the study, we want to clarify three issues: Is sonification applicable in the context of time-dependent medical data? Was the mapping we chose appropriate? Does the mapping help to give users an overall impression of the development of the patients?

We tested six subjects (two therapists, two musicians, two laypersons). Five of the subjects were male and one was female. The age range was between 25 and 55 years. The evaluation methods were observation, a semi-structured interview after the test and the analysis of the diagrams the subjects had to draw (see below). The testing consisted of an introduction, a learning phase and a test phase. The introduction consisted of an audiometric test and a short introduction into the domain. During the test phase, they had to assess the overall development of six patients based on the values of the four dimensions described above. For most of the dimensions, values for five points in time were available. They had to decide quickly whether the patients' development was positive or negative. Then they had to draw a curve of the development of these patients (see Figure 2). During the tests the subjects were allowed to listen to the reference sounds as often as they wanted. They did this quite frequently, some of them more than 100 times. It took the subjects between six (musician) and sixteen minutes to finish all the tasks. The whole user test took approximately one hour, including the audiometric test, the learning phase, the test phase and the interview.

There is some indication that the sonification system is relatively successful. The evaluation shows that the subjects mostly stick to one specific way of forming an overall impression of the patients' development. The subjects computed the average of the single values of the single dimensions (resulting from the questionnaires). Figure 1a represents an "ideal" diagram of one patient based on the numerical results from the questionnaires. In approximately 80% of the cases the subjects adopted this approach. The musicians were in general more accurate than the other subjects. The diagrams (see Figure 1b) indicate the overall development of the patients. The overall form of the curves is fairly correct, but the absolute level of the subjects' assessments does not conform to the reference curve. It is an open question whether this can be improved by practice.

The observation of the subjects indicates that they could describe the single values fairly well and understood the mapping. Differences in the quality of the mappings could be observed. Subjects complained that especially Beck's Depression Inventory was difficult to perceive, especially because of the use of

volume as a method to represent the value of the variable. The mapping of self-efficacy (rhythm), on the other hand, could be perceived very well. Most subjects started with this variable when trying to give their assessments.

Observation and interview yielded additional information. In general, the mappings were found to be quite good (apart from the Beck's Depression Index). One question asked in the interview was whether the subjects tended to interpret the sound as a whole or rather tried to assess every single component separately. There does not seem to be a consistent strategy. Some subjects adopted a more holistic approach, others did not. Observation indicates that concentration seems to be a very important factor in sonification. Some of the subjects remarked that using sound as a representation of quantitative values was quite unusual which made it so difficult. It is an open question whether this is a kind of literacy which can be learned or not.

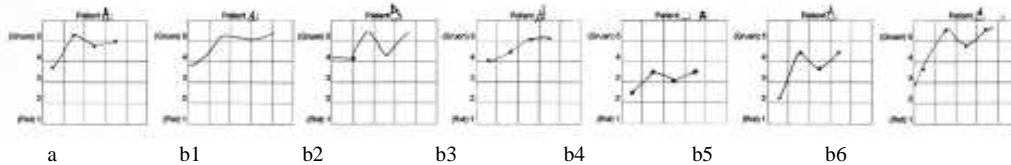


Figure 1. a. Reference value, b.1-6: Subjects' assessments of patient A's development

## 4. CONCLUSION

In this paper, we presented results from a pilot study investigating the usefulness of adopting sonification for the analysis of time-dependent data in medicine. The results seem to be promising, and a more in-depth study is certainly necessary. In answer to our questions our results show that sonification is, in principle, applicable in this context, although subjects were certainly not used to this kind of presentation of information. To analyze sonifications is certainly taxing. This might be due to the fact that this is an unusual form of presentation. Further research adopting longer phases of learning and adapting to the system is necessary to clarify this issue. The mapping seems to be appropriate.

Subjects could identify most of the sounds we presented to them, although some modifications are certainly necessary. We did not get a definitive answer to the question whether our mapping enables users to get an overall impression of the patients' development. Subjects seem to adopt different strategies. More research is necessary to clarify this issue. It was especially valuable to have potential users (the two therapists) and musicians as subjects. In the future, we want to do additional research using more data and to combine the sonification with the already existing visualization.

## REFERENCES

- Brewster, S., 2002. Nonspeech auditory output. In Jacko, J.A., Sears, S. (eds.) *Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications*, Lawrence Erlbaum, Mahwah, NJ, 220-239.
- Brown, L.M., Brewster, S.A., Ramloll, S.A., Burton, R., Riedel, B., 2003. Design guidelines for audio presentation of graphs and tables. In Brazil, E., Shinn-Cunningham, B. (eds.) *Proc. ICAD 2003*, 284-287.
- de Campo, A., Dayé, C., Frauenberger, C., Vogt, K., Wallisch, A., Eckel, G., 2006. Sonification as an interdisciplinary working process. In *Proc. ICAD 2006*.
- Flowers, J.H., 2005. Thirteen years of reflection on auditory graphing: promises, pitfalls, and potential new directions. In *Proc. ICAD 2005*.
- Rester, M., Pohl, M., Wiltner, S., Hinum, K., Miksch, S., Popow, C., Ohmann, S., 2007. Evaluating an InfoVis Technique Using Insight Reports. In: *Proc. Information Visualization Conference 2007 (IV '07)*, S. 693-700
- Salvador, V., Minghim, R., Levkowitz, H., 2005. User evaluations of interactive multimodal data presentations. In *Proc. IV' 05*, 11-16.
- Walker, B.N., Mauney, L.M., 2010. Universal design of auditory graphs: a comparison of sonification mappings for visually impaired and sighted listeners. *ACM Transactions on Accessible Computing* 2, 3.
- Zhao, H., Plaisant, C., Shneiderman, B., 2008. Data sonification for users with visual impairment. *ACM Transactions on Computer-Human Interaction* 15, 1.

# VISUALIZING THE IMPACT OF CONTENT-BASED SIMILARITY AND SPATIAL DISTANCE ON BOOK RECOMMENDATIONS

Sönke Knoch and Alexander Kröner  
*German Research Center for Artificial Intelligence  
Saarbrücken, Germany*

## ABSTRACT

Starting from the assumption that an integration of digital services might help brick-and-mortar bookstores to compete with online sellers, we propose a recommender system which takes into account not only content-based similarity of books, but also local availability and spatial distance. This article addresses work in progress; its focus is on a particular question tightly connected to the user interface design of such a system: how to communicate results from these different information sources to a customer in a transparent way. We contribute to this topic with several design proposals and a discussion of feedback collected in a user study.

## KEYWORDS

Keywords-Recommendations, spatial information, transparency, bookstore, requirements.

## 1. INTRODUCTION

Nowadays, bookselling basically exists in two main forms – physically in a bookstore, or virtually online. While both forms seem to co-exist, the growing success of eBooks might become a serious challenge for brick-and-mortar bookstores [1]. The latter ones may address this challenge by combining the advantage of the books' physical presence with features known from online retailers. For instance, kiosk systems such as [2] allow a store's customer to search for a book, check if it is in stock, and get a print-out of the particulars of a book or a search. However, actually retrieving a book in such an extended store nevertheless has a very physical aspect: customers have to walk in order to reach a book of interest – a potential burden if, e.g., several books need to be investigated. A digital assistant could address this issue by recommending books which are not only of interest for, but also close to the user. The basis of such a service is the combination of a book's content-related features with the given environment's spatial features, such as locations of shelves, books, and supporting infrastructure. This article focuses on the presentation aspect of this issue – how to make the impact of both feature categories on recommendations transparent to the user. In the following, after a brief review of related work, we describe a kiosk prototype which recommends books on the basis of content-based similarity as well as spatial distance. For the presentation of recommendations, various design alternatives are proposed. Then, a small-scale study aiming at feedback concerning these alternatives is discussed. The article concludes with a discussion of the results and their impact on future work.

## 2. RELATED WORK

Our approach to interaction aims at guiding the user from a fixed kiosk to books and shelves of interest. It could be combined with systems which support the user on the way to or at the shelf – e.g., an intelligent environment able to highlight relevant areas in the user's surrounding (e.g., [3]), or a mobile augmented reality system for book spine recognition which allows retrieving information about books on a bookshelf by snapping a photo (see, e.g., [4]). Regarding the proposed recommendation approach, a study of Ducheneaut et al. sheds light on benefits from combining collaborative filtering with other information sources, including

spatial distance [5]. It emphasizes the need to make the impact of these sources transparent and customizable. For presenting such recommendations, various approaches are known from research concerning location-based user support in mobile scenarios, especially if decisions about people, locations, and artifacts around the user are required. For instance, the SocialSearchBrowser exploits the user's social network in order to enhance a (mobile) map-based display of queries related to the user's location [6]. Our experiment includes a similar option; however, we were also interested in visualizations providing an abstraction from the environment's spatial details. This is related to visual information seeking for literature, e.g., in the area of digital libraries [7]. Here, our special interest was to convey a combination of spatial and content-related features in an easy-to-understand and transparent way.

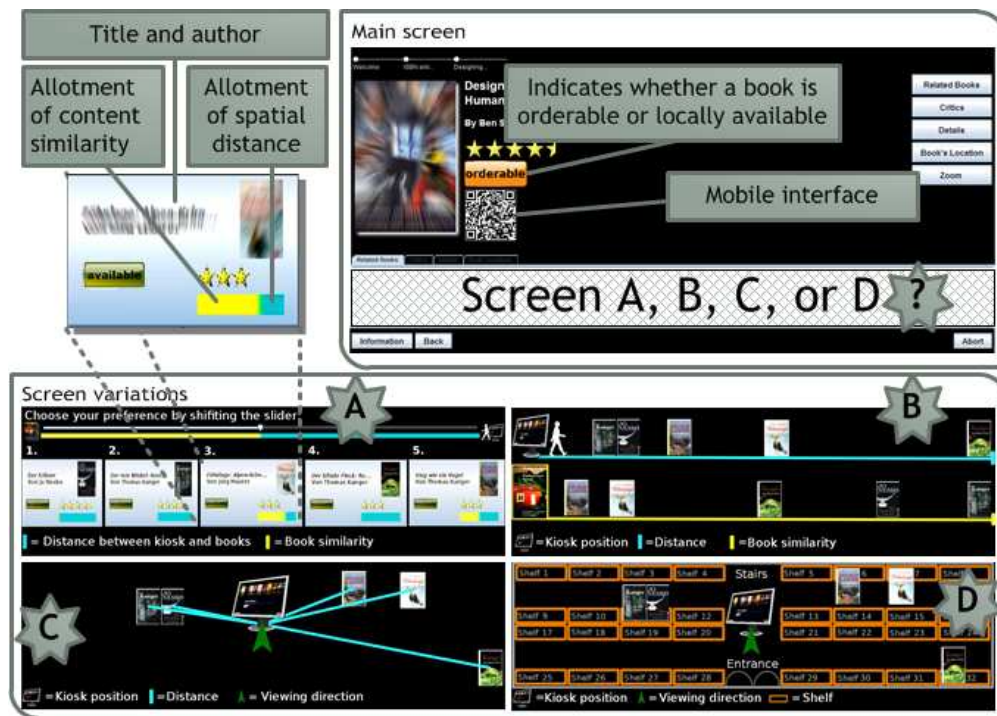


Figure 1. The kiosk's user interface and screen variations A, B, C, D for the display of recommendations

### 3. SYSTEM PROTOTYPE

*Infrastructure.* Our prototype system relies on books tagged with visual codes or radio-frequency identification (RFID), and a virtual model of a store's physical structure. A relational database maintains coordinates for each object in the store. The shelves are subdivided into sections to allow sub categorization in the store. The locally available books are arranged into sections. In this early prototype, the actual assignment of coordinates to objects is done manually.

*Interaction.* The conceptual model of the interaction between user and kiosk follows the idea of direct manipulation. The user interacts via a touch screen and a code reader. Via the latter one, the customer may use a book at hand to start searching and browsing the shop's inventory. Alternatively, he or she may start an inquiry from an entrance screen displaying popular or new books.

*Recommender.* If the user requests similar books for a book at hand, the system computes a ranked list in two steps: (1) it determines the content-based similarity of books available in the store using the recommendations included in Amazon Web Services (AWS) [8], and (2) the resulting list of books is re-ranked using information about their spatial distance from the kiosk. Since the full book content was not available at the time of implementation, we decided to use the book's sales rank to measure similarity among books. In this process, the spatial distance is calculated by the Euclidean metric. Both factors are weighted accordingly to their importance as expressed by the user through a slider (default configuration: 0.5=neutral).

*Visualization.* For presenting recommendations, a prototype was implemented which stays close to common list-based presentations (see Figure 1, screen A). Feedback of 5 users (IT experts) indicated a need for more transparency concerning the impact of spatial data. In order to address this request, we created mock-ups of three additional screens B, C, and D which might either replace or complement the original list-based display.

Screen A provides an overview of recommended books in form of a single, cumulated list in descending order concerning relevance. To provide insight to the impact of spatial and content distance per book, a chart-like rectangle shows the allotment by painting the respective factor's color. A slider allows the user to adjust the weighting between both factors. Shifting the slider changes the ordering of the books.

Screen B basically presents the recommender's result before content and spatial distances are combined. It relies on two separate fixed lists, one for spatial distance to the kiosk, one for similarity of content to the book at hand.

Screen C conveys information about the direction the customer would have to walk to obtain the recommended books. It uses a star-like graph with books as leafs and the kiosk as source, and indicates the viewing direction.

Screen D aims at indoor navigation support. It relies on an abstract map-based display of the store's entrance, stairs and shelves in form of rectangles together with the kiosk location. The shelves are numbered and books are faded in above the shelf they belong to.

In summary, from A to D these screens provide an increasing accuracy of spatial information. The 1-D screens A and B show a simple (spatial and content) distance measure, whereas the 2-D screens C and D come up with information on product direction and location. However, the content-based similarity measure tends to disappear from A to D, while the spatial information is stated more precisely.

## 4. EXPERIMENT

To verify the appropriateness of these screens for the given task – and eventually to identify the most appropriate combination of screens –, we conducted an experiment with a particular focus on answers to the following questions:

**Organization.** What opinion has a customer on spatial distance vs. what kind of representation does he or she prefer, accumulated variable list A, split fixed list B, graph C or map D?

**Distance Visualization.** Is a customer confident with a simple 1-D spatial distance measure or does he or she need a 2-D visualization which additionally informs about product locations?

**Classification.** Is it possible to identify behavioral patterns that can be used to classify users according to their spatial preferences for the purpose of personalization?

To find the answers, a seven-page questionnaire was used in a 45 minutes survey, which was divided into three phases. The paper-based questionnaires comprised closed and open questions. For closed questions check boxes and ranges in form of a five level Likert scale were chosen to gather the data. A supervisor guided the participants through the experiment.

### 4.1 Sample and Procedure

The study was conducted at a university. Altogether 21 people (11 *f*, 10 *m*) with an average age of 27 (median = 25) took part in the experiment. All participants were internet users and most of them (81%) *buy online multiple times per year, once a month or multiple times per month.*

Phase 1 started with a short introduction in text form. The participants were asked to imagine a bookstore where an interactive kiosk system equipped with a touch screen can be found; the screen would represent the status after a book has been selected through browsing or code scanning. Then, the participants had to fill out a first questionnaire containing general questions and questions on the customer's intention before and behavior during a bookstore visit.

In phase 2, the screens were presented on a 15.4" laptop screen in the fixed order A, B, C, D. To check the participant's ability to interpret the screens, the screen-related screen contained questions about layout and functionality. Then, it focused on the meaning of the presented recommendation and the comparison of B, C, D with A. After the last screen, the participants were asked to declare the most and second most helpful



screen. Additionally, we asked for the preferred screen under three different behavioral patterns: *to buy one single book*, *multiple books* or *spontaneously*. At this point, the demographic data were gathered.

In phase 3, for each of the screens B, C, D, the participants had to plan their favorite sequence of 3 recommended books on a route with a kiosk-to-kiosk connection. A corresponding questionnaire included numbered pictures of 3 book covers. The display comprised 5 books, including all of the 3 books in question. The experiment was conducted separately for each screen; the books were always the same. A supervisor measured the time the participants needed to define a sequence he or she found ideal to reach the books. The background for this test was that we wanted to test for B if the user chooses the correct distance list and for C and D if the abstract graph outranks the map.

## 5. RESULT AND DISCUSSION

In the following, the feedback on the three questions concerning *Organization*, *Distance Visualization*, and *Classification* is presented.

### 5.1 Feedback related to Organization

Asked about the personal importance ascribed to the spatial distance, the result was nearly balanced: 48% found that spatial distance is important. In contrast, according to answers to the question for the most helpful screen, the map-based representation D performed best, followed by B and A (see Table 1). In the route planning task, screen C performs best with an average planning duration of 14.18 sec, followed by screen D with 15.74 sec, and screen B with 18.92 sec.

Table 1. Screen rating, the percentage of participants that chose two screens as 1st and 2nd choice. To indicate differences in choice between the genders, the result is split according to the participants' sex.

Rank	1st choice	2nd choice	Rating	Female	Male
1	Screen D	Screen B/C	65%	0%	100%
2	Screen B	Screen D	25%	80%	20%
3	Screen A	Screen D	10%	54%	46%

Barely half of the participants think that spatial distance is important, although the sample consists predominantly of online buying young people. This result indicates that people ascribe an importance to the distance between them and the desired product. On the other side, a list-based representation seems not to be the desired one: the graph and map representations beat the other screens, which was shown through the rating task. As expected, the route planning task was dominated by the information poorest screen C, closely followed by D. Finally, we can conclude that the participants ascribe an importance to spatial distance and the screen with the highest degree of spatial information performs best. The fact that on rank 3 A and D occur with the best balance between genders is an indicator for a good combination.

### 5.2 Feedback related to Distance Visualization

*Screen A.* We wanted to check if the meaning of the slider is transparent for a user. The participants were asked about their assumptions concerning the effect of a shift to the outer left. 43% replied correctly that the ordering of the presented books would change. 38% were also able to identify the kind of change in which direction similar books move regarding content and regarding physical distance. Only 19% answered totally correct. Interestingly, 24% thought the most content-related books instead of the closest books would slide on top of the list. They changed the meaning of the two colors. 47% identified the 3rd book as the most appropriated book instead of the 1st book. When we asked for the factors which are used to calculate the ranking, 33% were able to identify them.

*Screen B.* Regarding the list-like distance visualization, 48% thought that they could make a clear statement on product locations, which is an impossible task. 86% interpreted the leftmost and rightmost books as *spatially most distant*, which is false. The content-based distance of the same books in comparison

with their spatial distance was seen as relatively smaller by 76%. Solving the route planning task, 81% simply planned the route from the left to the right as expected and used the correct list.

*Screen C.* 10% saw a content-based relationship between two of the recommended books, which was not intended. Vice versa only 19% saw no such relationship. The remaining 71% checked no box. The location was perceived well by 91% of the participants. Interestingly, 57% of the people thought that the two books in the upper right corner were close to the kiosk – despite the lack of a distance unit. 43% thought that both were located in the same shelf row.

*Screen D.* The books were located as accurate as possible by nearly all participants correctly. At the route planning task, 86% planned to visit the two books in the same book shelf corridor successively.

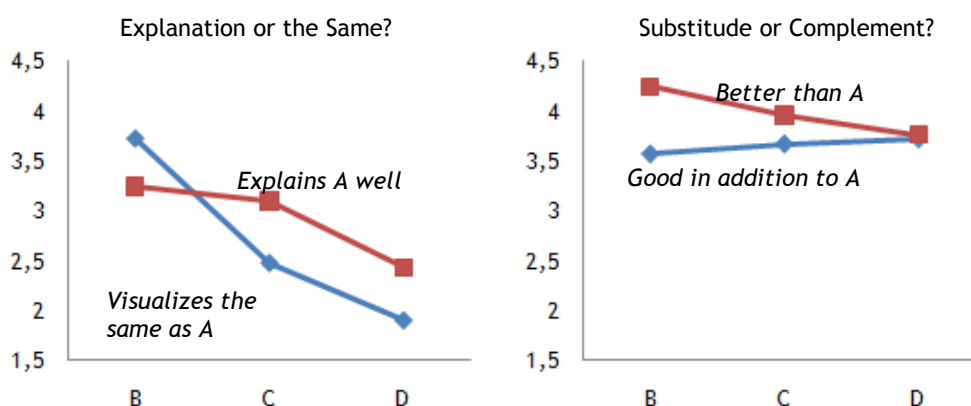


Figure 2. Screens B, C, D compared to A on a scale between 1=strongly disagree and 5=strongly agree.

*Screen comparison:* Figure 2 shows on the left that the screens are explaining A decreasing from B to D, while they visualize strictly decreasing the *same as A*. The diagram on the right shows that all screens are perceived as *better than A* and *good in addition to A*. From B to D, there is a tendency that the screens fit increasingly *better in addition to A* (complement). In parallel, the screens are decreasingly *better than A* (substitute).

The feedback on 1-D screen A shows that a cumulated list representation with a slider to edit preferences and the current color coding is non-transparent to the user. The split list in B animates the user to interpret not intended coherences between the books, because they are connected through a colored line. The star-like graph in C is easy to understand, but some people tend to interpret coherences like content-based relationships, closeness and shelf locations although there is no connecting line. The map screen D bridges this gap and is well understood by a high majority of the participants. The screen comparison results that A and D complement one another best, which strengthens the indicator on this combination from the feedback on organization. Finally, we can conclude that the user has problems to understand the distance measure in a 1-D visualization. A 2-D visualization leads to less confusion and to a more precise interpretation. The map is the preferred visualization, because the least misunderstandings occurred, it is seen as best in addition and least the same as A (complement).

### 5.3 Feedback related to Classification

The sample can be divided into online (62%) and offline buyers (38%). Many participants are goal-oriented customers: 57% investigate online before they visit a bookstore, 67% commonly visit a bookstore because of one single book and even 19% prepare a book list. Regarding shopping behavior, the majority of customers in our sample commonly just browse a bookstore (91%). At the end of a bookstore visit, 76% commonly buy one single book and 30% buy multiple books; multiple answers were possible. So far, two groups exist concerning buy behavior: *single* and *multi book buyers*. Many people buy more books than they plan to. Thus, a group of people who buy *spontaneously* exists. The question of the most suitable screen under three different conditions, the behavioural pattern, resulted in the dependency of the demand on different behavioral patterns of the bookstore customer. Table 2 shows that single book buyers prefer D, multi book buyers are indifferent between B and D and spontaneous ones prefer B and refuse D.

Table 2. Screen preferences under three conditions.

Condition: behavioural pattern	A	B	C	D
Buy one single book	15%	0%	10%	75%
Buy multiple books	19%	38%	10%	33%
Spontaneous	33%	48%	19%	0%

## 6. CONCLUSION

This article addressed an approach to integrate spatial features of books into a book recommendation system which could be deployed in a brick-and-mortar bookstore. An experiment addressed various ways to visualize recommendations computed by such a system. Results from that experiment indicate that spatial information actually might help the customers of a bookstore to plan their shopping trip. Users benefited most from the integration of spatial data if visualization supported route planning inside the store. Furthermore, customers' preferences concerning spatial data differed which could be exploited for classifying users according to their spatial preferences.

These observations have several limitations, which might become subject of future work. For instance, participants didn't experience any negative feedback after bad route planning which might have affected their perception of spatial support. Therefore, an experiment is needed where participants actually have to perform tasks similar to the reported ones in a store-like scenario with efforts needed to reach a book. Another limitation is the specific (narrow) user group in this experiment; here further experiments with other user groups (e. g., elderly people) might provide additional insights into the topic. Future work will have to address the potential and the limitations of these observations. Finally, a technical extension of the prototype could provide a visualization which adapts to behavioural patterns and thus spatial preferences of customers.

## ACKNOWLEDGEMENT

This work was developed in the context of the project *Allianz Digitaler Warenfluss* (ADiWa) that is funded by the German Federal Ministry of Education and Research under grant 01IA08006.

## REFERENCES

- [1] International Digital Publishing Forum. Wholesale eBook sales statistics. [http://idpf.org/doc\\_library/industrystats.htm](http://idpf.org/doc_library/industrystats.htm), accessed on August 10, 2010.
- [2] Phosphor Essence, "Book Finder Kiosk," <http://bookfinderkiosk.com/>, accessed on August 18, 2010.
- [3] Butz, A., Schneider, M., and Spassova, L., 2004. SearchLight - A Lightweight Search Function for Pervasive Environments. *Proceedings of the Second International Conference on Pervasive Computing*, LNCS 3001, pp. 351-356. Springer: Berlin, Heidelberg.
- [4] Chen, D., Tsai, S., Hsu, C.-H., Singh, J. P., and Girod, B., 2011. Mobile augmented reality for books on a shelf. *Proceedings of IEEE Workshop on Visual Content Identification and Search (VCIDS)*, to appear.
- [5] Ducheneaut, N., Partridge, K., Huang, Q., Price, B., Roberts, M., Chi, E.H., Bellotti, V., and Begole B., 2009. Collaborative Filtering Is Not Enough? Experiments with a Mixed-Model Recommender for Leisure Activities. *Proceedings of the 17th International Conference on User Modeling, Adaptation and Personalization (UMAP 2009)*, LNCS 5535, pp. 295-306. Springer: Berlin, Heidelberg.
- [6] Church, K., Neumann, J., Cherubini, M., and Oliver, N., 2010. SocialSearchBrowser: A novel mobile search and information discovery tool. *Proceedings of the 15th international conference on Intelligent user interfaces (IUI 2010)*, pp. 101-110. ACM Press, New York, NY, USA.
- [7] Heilig, M., Demarmels, M., König, W.A., Gerken, J., Rexhausen, S., Jetter, H.C., and Reiterer, H., 2008. MedioVis: visual information seeking in digital libraries. *Proceedings of the working conference on Advanced visual interfaces (AVI 2008)*, pp. 490-491. ACM Press, New York, NY, USA.
- [8] Amazon.com. Amazon Web Services. <http://aws.amazon.com/>, accessed on September 1, 2010.

# SELECT'N'LABEL – DYNAMIC BOUNDARY LABEL PLACEMENT FOR MOBILE TOUCH DEVICES

Dominik Magin and Daniel Steffen

*German Research Center for Artificial Intelligence (DFKI) GmbH  
Kaiserslautern, Germany*

## ABSTRACT

Including 3G, long battery life, and touch screens, mobile devices are about to replace notebooks and desktop computers. In professional practice, such as map-based large area maintenance jobs, mobile devices allow users to have only one device with all necessary information. However, the usually lower performance of such devices, reduced display size, and touch interaction limit the amount of information to be displayed.

This paper introduces the *SELECT'N'LABEL* technique: a dynamic label selection approach to interact with a labeled map and to select labels in an area of interest illustrated with a fisheye magnification. Furthermore, we introduce a dedicated 3-sided boundary label placement algorithm for mobile touch devices. Here, we consider performance, display size, and interaction mode.

## KEYWORDS

Boundary Label Placement, Mobile Touch Devices, Fisheye

## 1. INTRODUCTION

Map labeling on mobile devices is a pretty new application area. To work and interact with mobile touch devices different factors that limit user interaction have to be considered. In contrast to standard computer monitors, display size and resolution of mobile devices are much smaller and cannot display the same amount of information. Furthermore, the interaction via a touchscreen is slightly different to mouse and keyboard interaction. Firstly, a precise interaction depends both on the size of the interaction area and on the size of the user's fingertips. Thus, the interaction elements need to have a minimum size in order to avoid false interaction. Secondly, while interacting with the device, it is possible that the user's hand, finger or arm hides parts of the screen.

This paper describes a solution of a performance label placement algorithm and how we deal with the limited screen-size. We implemented a fisheye-view to magnify an area of interest and select labels. The labels are placed by our algorithm in an optimal position concerning minimum leader length.

## 2. RELATED WORK

Much research has been conducted on the subject of label placement. Azuma (2003) distinguishes between internal and external label placement. In case of an internal label, the label is placed inside the object it describes. For an external label placement the annotation is placed outside the object. Baudisch (2008) states, that *"the use of internal labels can reduce visual search as targets and label are associated by proximity, while users need to trace a line in order to locate an external label"*. However, in the conducted user study they consider a Tablet PC with a pen as input device.

Similar to internal label placement, point feature labeling (do Nascimento (2003)) places the labels next to an object. However, point feature labeling on mobile devices has potential limitations. Considering many feature points on a map it is difficult to place multiple labels within a small distance to each other. In contrast to normal sized displays, a map must be reduced in order to fit on the mobile device's screen. Not only the map becomes smaller, also the points on the map get closer. Additionally, the size of the labels cannot be

scaled proportionally to scaling of the map, since they must have a minimum size for a proper interaction. This results in the fact that label overlapping could occur (cf. Figure 1a).

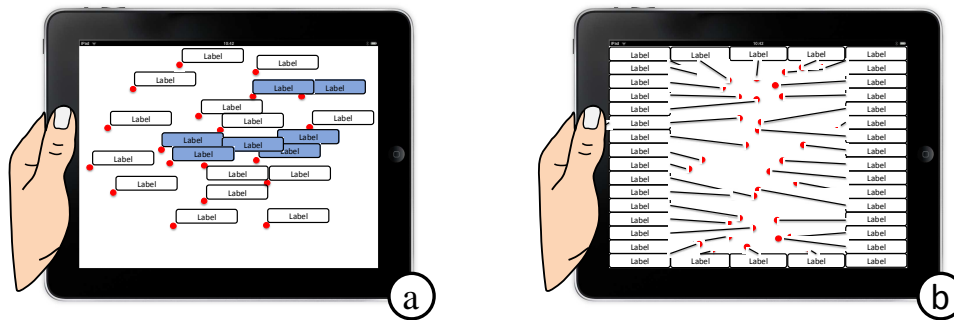


Figure 1. Problems of labeling techniques: While point feature labeling introduces overlapping (a), boundary label placement results in clutter caused by the leaders (b).

Placing labels on the screen boundaries (Bekos (2007)) causes clutter (cf. Figure 1b) because of the lines that lead from the feature point to the label. Moreover, the user needs to trace the lines of a label. The small display size is advantageous, because the line-length is smaller than on a normal-sized display. Thus, the tracing time and distortion by the user is reduced. Using a suitable label sorting technique can reduce the clutter to a minimum. Therefore, we consider boundary label placement in the following.

Fuchs et al. (2006) consider an adaptive labeling approach and the restricted resources on a mobile device. The calculation of the label positions, however, is done on a distributed computation platform. So far no literature is known that deals with dynamic boundary labeling on mobile touch devices.

Bekos et al. (2007) consider different kinds of boundary labeling, like labels of identical or different size, straight-line or rectilinear leaders and 1-/2-/4-sided boundary labeling. Additionally, Nöllenburg et al. (2010) describe a one-sided clustered labeling approach. Here, the labels are placed in an optimal position according to the minimum leader length. But they only consider boundary labeling for one display side.

This paper tackles the above-mentioned problems. In the following section we describe – based on Nöllenburg’s idea – a 3-sided boundary label placement. Finally, we will present a solution on how to reduce the amount of displayed labels and interact with labels on a map.

### 3. DYNAMIC BOUNDARY LABEL PLACEMENT

In this section we describe a dynamic boundary label placement technique. The keyword dynamic means that only a group of labels is displayed and the arrangement of the labels is computed *dynamically*. In section 4 we present a label selection technique to select a group of labels.

Nöllenburg’s 1-sided clustered labeling approach allows to group labels dynamically, introduces gaps between label groups and calculates the best label position according to minimum leader length (cf. Figure 2a). With only a few labels, Nöllenburg’s algorithm produces very good results. However, the whole position calculation with all its cluster merging and up- and downward passes needs too much time and is not suitable for a mobile device. Secondly, we deal with a high amount of labels. Most of the time, clustered labeling calculated with Nöllenburg’s algorithm results in only one or two clusters. For our application we never reached more than two clusters. For a huge amount of labels the number of clusters converges against one (cf. Figure 2c). Regarding the above mentioned we simplified Nöllenburg’s approach. Instead of using multiple clusters, we only use one cluster.

In our approach we calculate the position of one cluster  $C$  as following: we determine the  $y$ -position  $y_1$  of the point nearest to the bottom and the  $y$ -position  $y_2$  of the point nearest to the top. Afterwards, we calculate the midpoint of these two positions,  $m = (y_1 + (y_2 - y_1)/2)$ .

The cluster  $C$  has the extension  $e = label\_amount * label\_height$ . Now we consider  $m$  as the center of the cluster  $C$ . From bottom to top the cluster  $C$  starts at  $m - e/2$  and ends at  $m + e/2$ . In this cluster the labels are placed above each other towards their  $y$ -position (cf. Figure 2b). Compared to Nöllenburg this algorithm provides most of the time the same results, but is a lot faster (cf. Figure 2c).

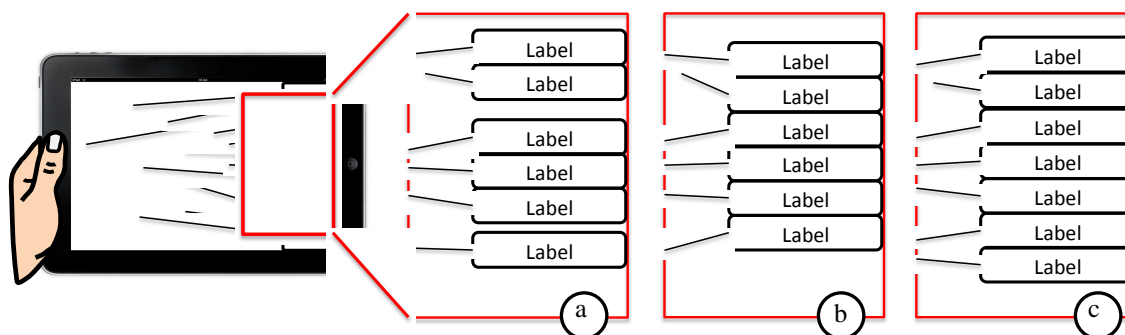


Figure 2. Nöllenburg's labeling approach (a) allows gaps between label groups whereas our algorithm (b) treats the labels as one group. However, when introducing one more label (c) both methods produce the same result.

Nöllenburg et al. (2010) consider 1-sided, Bekos et al. (2007) consider 1-/2-/4-sided labeling but to our knowledge no approach deals with 3-sided boundary labeling. The advantage of 2-/3-/4-sided labeling is, that we can place much more labels on the screen than on one side. In our approach we implemented a 3-sided boundary label placement for top, left and right. The purpose we do not consider the bottom area for labels is, that when interacting with the map (cf. section 4) the user in general hides the bottom area with his finger, hand, or arm (cf. Figure 4).

For the 3-sided boundary labeling, we create three groups of labels/points. Each group belongs to one of the boundaries (left, right or top). We use hierarchical clustering to assign the labels/points to the three groups. Hierarchical clustering allows to group points according to their distance to each other. Initially, every point forms a single cluster. As long as the number of clusters exceeds the amount of three, the two clusters with least distance to each other are merged (bottom-up approach). Using the Euclidean distance we define the minimum distance between two clusters as the distance between the two nearest points of each cluster. For these three groups we determine the position (center) to each other in order to allocate them to the boundaries. For each boundary-side we get one group of labels that will be displayed. To each of these three groups our algorithm is applied (for the top group, we use the x-positions, instead of y).

## 4. INTERACTION TECHNIQUES

According to Shneiderman (1996) a visualization should provide both an overview and details on demand. Due to the small screen sizes, the information density of a given presentation has to be carefully managed for the visualization to remain effective. Non-linear distortion techniques enable comprehensive detail-in-overview representations of large images on small screens (cf. Figure 3). Furnas (1986) Generalized Fisheye View allows, even on mobile devices, to display a map as *overview* and still selecting an area of that map as *details-on-demand*. For the fact that we cannot display all labels at once, we need a solution to select the labels of interest. All points of the labels are visualized on the map, while the corresponding labels are hidden. We use Furnas' fisheye-view not only for getting a detailed view of an area of the map, but also for selecting the labels we want to see. The points in Figure 4 are a way to small to press, especially when they are very close to each other. Instead of pushing a point, you select the area you want to magnify and the desired labels are shown (cf. Figure 4). We implemented a feature for scrolling through the datasets. By tapping – hold-down and moving your finger – the fisheye follows the finger and hides the labels of the previous areas and displays the labels of the current magnified area. As a result, the labels inside the magnified area are handed over to our dynamic boundary labeling (cf. Figure 4). Combining these two techniques we have the means to display and interact with large datasets on mobile devices. Using the fisheye-view we introduced a technique that emphasizes the area currently being in the focus, and selects the labels, which are displayed. The labels are ordered by the dynamic boundary labeling described in section 2.

## 5. CONCLUSION AND FUTURE WORK

In this paper we presented a three-sided boundary label placement method that is qualified for the use on a mobile device. In combination with the presented fisheye-view the user has the possibility to scroll easily through a huge amount of datasets without getting distracted by them. Our approach allows more points on the map, as there is space for labels; especially on small screens this is an advantage.

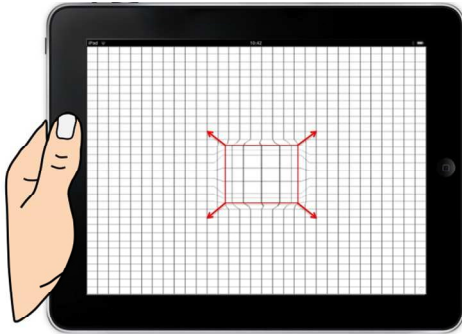


Figure 3. An area of the screen is magnified while the borders are distorted in order to fit the screen.

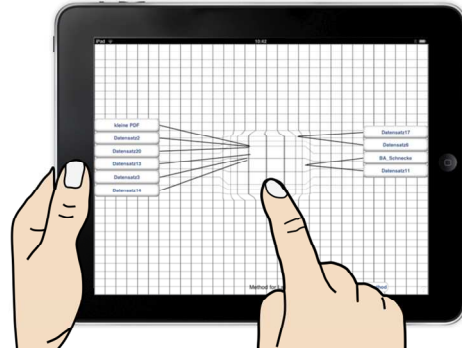


Figure 4. Tapping on the screen, an area is magnified and the labels inside that are displayed.

Until now we did not consider the case that the amount of labels in the fisheye-area is too big. Methods need to be figured out, how to scroll through these labels. Moreover, a user study will be conducted in order to clarify, if the advantages of boundary labeling really outbalance. Finally, we want to determine the best size and magnification factor for the fisheye view in order to gain user satisfaction.

## ACKNOWLEDGEMENT

This research was supported by the German Federal Ministry of Education and Research (BMBF) within the project SYNERGIE (Grant No. 02WQ1102).

## REFERENCES

- Azuma, R. & Furmanski, C. (2003), Evaluating Label Placement for Augmented Reality View Management, *in* 'Proceedings of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality', IEEE Computer Society, Washington, DC, USA, pp. 66.
- Baudisch, P.; Zotov, A.; Cutrell, E. & Hinckley, K. (2008), Starburst: a target expansion algorithm for non-uniform target distributions, *in* 'Proceedings of the working conference on Advanced visual interfaces', ACM, New York, NY, USA
- Bekos, M. A.; Kaufmann, M.; Symvonis, A. & Wolff, A. (2007), 'Boundary labeling: Models and efficient algorithms for rectangular maps', *Comput. Geom. Theory Appl.* **36**, 215--236.
- Fuchs, G.; Luboschik, M.; Hartmann, K.; Ali, K.; Schumann, H. & Strothotte, T. (2006), Adaptive Labeling for Interactive Mobile Information Systems, *in* 'Information Visualization, 2006. IV 2006. Tenth International Conference on', pp. 453 -459.
- Furnas, G. W. (1986), Generalized fisheye views, *in* 'Proceedings of the SIGCHI conference on Human factors in computing systems', ACM, New York, NY, USA, pp. 16--23.
- Nöllenburg, M.; Polishchuk, V. & Sysikaski, M. (2010), Dynamic one-sided boundary labeling, *in* 'Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems', ACM, New York, NY, USA, pp. 310--319.
- do Nascimento, H. A. D. & Eades, P. (2003), User hints for map labelling, *in* 'Proceedings of the 26th Australasian computer science conference - Volume 16', Australian Computer Society, Inc., Darlinghurst, Australia, Australia.
- Shneiderman, B. (1996), The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations, *in* 'Proceedings of the 1996 IEEE Symposium on Visual Languages', IEEE Computer Society, Washington, DC, USA.

# TANGIBILITY AND HUMAN-COMPUTER INTERACTION: AN ALTERNATIVE APPROACH TO AFFORDANCE

Jean Caelen, Mélanie Becker and Anne Pellegrin  
*Multicom, Laboratoire d'Informatique de Grenoble,  
UMR 5217, BP 53, F-38041 Grenoble cedex 9, France*

## ABSTRACT

This paper deals with a tangible interface, Tangisense, and aims at presenting the contribution of tangibility to the field of man-machine interaction from the point of view of affordance. Some experiments build on a tangible table Tangisense, developed by the laboratory, demonstrate that the affordance plays different roles related to the context, the type of task and the collaborative interaction with others users. We show that the affordance of objects is a complex concept that depends not only on the appearance of the object but also on what users do in a collaborative context.

## KEYWORDS

Tangibility, Interactive table, Affordance, Collaborative work.

## 1. INTRODUCTION

For 25 years, tangible interfaces (TUI = Tangible User Interfaces) have been spreading more and more (Blackwell et al., 2007), so that it could eventually replace the traditional graphical user interface of applications such as games, participatory design work, the mock-up design, etc. (Kubicki et al., 2009). The definition by Ishii and Ullmer (1997) is very concise but easily understandable: it comes to manipulating real objects that are integrated in a virtual environment and have digital capabilities of interaction – it is not just to interact with virtual objects on a screen but more interacting directly by and on such objects. Thus, thanks to the TUI, you don't have to insert an artifact of communication in the sensorimotor process that is established between the user and the environment or to break the continuity of touch created by the mouse as in graphical user interfaces (Moggridge, 2006). In this type of tangible interface, the object is both a tool of the interface, object of the application and device for the interaction, which raises its ambivalent status. Unlike the interaction with a computer screen where the user must manipulate the mouse, tangible interfaces require different postures for the body. The representations are outsourced; the feedback may be visual, tactile or haptic (Hornecker and Buur, 2006).

In this paper, we examine the tangibility in terms of affordance and collaborative work from experiments with the interactive table Tangisense, exploiting the RFID technology (Radio Frequency Identification) and allowing collective interactions with tangible and virtual objects.

## 2. THE TANGISENSE TABLE AND TANGIBLE INTERFACES

### 2.1 General Characteristics

Tangisense is like a magnetic retina, able to detect and locate RFID tags pasted on various objects. The table consists in 25 blocks, each containing 64 antennas (8 x 8) by 2.5 cm square on a surface of 1m<sup>2</sup>. Each block contains a processor reading RFID antennas, an antenna multiplexer and a communication processor. The blocks are linked together by a control interface connected to the host computer by an Ethernet bus. With this technology, one can recognize tags for objects superimposed each on other and their positions in the



workspace. One can detect an object completely hidden under another bigger than himself, and find its position too. Each tag has a little memory capacity for storing information in the RFID.

## 2.2 The Table Objects

There are two different types of objects on the table that can be interactors: (a) the virtual objects (as in a GUI, such as scroll, buttons, etc...) and (b) tangible. The virtual objects are projected onto the table using an LCD situated on the surface of the table or a video projector placed vertically above the table. Tangible objects are equipped with RFID tags glued to their base. In these tags, it is possible to store information such as the latest movement, name of its owner, etc.

Several applications have been programmed for the table. In addition to a road traffic simulation, two applications were referred to musical production:

(a) Automaton Music: this is a software for interacting with instruments (tambourine, drums, horn, etc.). To do this, you must first lay on the table an object associated to a particular rhythm. Once installed, some LEDs will light up on the table. When an instrument is placed on the same area than the LEDs, the table will play the corresponding sound: drums or whatever.

(b) CD Judbox: the principle is similar to the foregoing; the difference lies in the type of object presented (sound textures instead of instruments). The purpose is to place an audio CD on the table that will extract the various sound textures. You may then register these textures through interaction objects and replay them at will. As indicated above, when an object is placed on the same area than the LEDs, the table will play the sound according to a specific rhythm.

## 2.3 Interaction, Posture and Collective Work

Tangible interfaces are a technology that allows the user to have a “natural” and “friendly” experience. Indeed, according to Ishii et al. (2001), these interfaces would give the feeling of being connected to the real world. Through the multisensory control that allows direct manipulation of objects, we have the opportunity to have richer expressive gestures that allow natural and intuitive interaction (Fiebrink and Morris, 2009). The advantages of this technology are numerous. Interaction with both hands is strongly encouraged, which is important because the physical capabilities of the hand and wrist are also rich (Fitzmaurice and al., 1995). What’s more, users can be placed anywhere around the table: entries are spatially and it significantly improves the ability to communicate with the computer. Physical artifacts facilitate interaction by making more direct interface, and especially more manipulable. This type of technology also encourages the collaboration of several people (Fitzmaurice and al., 1995). Finally, since the objects are manipulated in three dimensions on a horizontal surface, the vast majority of these tables offer a generous workspace and allow the user to act while talking and keeping an eye on what others are doing (Manches et al., 2009). One of the key points in this interaction remains the visibility of other users and, therefore, often the goal of understanding and gesture (Fiebrink and Morris, 2009). However, with this technology people fear to experience a physical collision or an encroach on the territory of a neighbor. In summary, tangible interfaces restore a significant role to manipulating and exploiting the dexterity acquired by humans in their daily environment (Couture et al., 2007). The interface improves multiplayer actions allowing better hand coordination, parallel actions, a better perception of space and changes in perspective by the possible moves around the table.

## 3. THE AFFORDANCE

Gibson introduced the concept of affordance in 1979 and made it a theoretical pillar of the ecological approach. The affordance of an object allows the user to perceived immediately how he may interact with this object. This concept has lead to numerous appropriations by different authors, especially by Norman (1999). According to Norman's, affordance corresponds to information needed in the world to act appropriately according to the project and the objectives of the concerned actor. The appearance of a device must therefore be able to provide relevant information on what possible actions are proposed by the item

(Norman, 1999). Mucchielli (2000) uses the term "holders of proposed interaction" to refer to objects affordances that have the property to offer some plausibility of their duties. With this in mind, it is mostly important for tangible objects to be affordant. Indeed, with TangiSense technology, tangible objects can only be perceived through their assigned functions. The principle of affordance is then running the risk of being lost. In summary, the affordance is a relationship established between the user and the objects in the environment through their properties but also according to the task to perform and in relation with others agents acting with these objects. The affordance problem has three dimensions: purpose, task, people group.

## 4. EXPERIMENT

To better understand the concept of affordance, an experiment was conducted in our laboratory (Becker, 2010) to measure the affordance of objects in a work situation with the interactive table TangiSense. The subjects were alone and had to perform a task with more or less affordable objects that they did not know (situation called direct) or had to perform these tasks with other people who knew how to manipulate objects (situation called indirect).

### 4.1 Research Hypothesis

We defined several assumptions:

- H1 - There are some differences of understanding between affordable or non affordable objects, regardless of the situation and players (in other words, the affordance belongs essentially to the nature of objects),

- H2 - The situation of interaction with an object affects its understanding and construction of affordance, in particular through the presence of other agents,

- H3 - The single or multi-functionality property of an object affects its understanding and the construction of affordance (the multi-functionality of an object is defined by the fact that it has a self affordance and an inherited affordance increased by its digital capabilities – for instance a glass with an interactive function will still be useful to drink).

The operating assumptions are that these differences are measurable, e.g. in terms of understanding time and ease of comprehension (number of errors in the subjective assessment of the task).

- H'1 – With a degree of equal affordance, interface objects will be understood more easily and faster than application objects, because of their cultural largest value (indeed people have acquired this culture of interface objects from recent and frequent use of the GUI).

- H'2 - For an affordable object the “direct situation” is more favorable to the comprehension and less favorable for the “indirect situation”.

- H'3 - The less affordable objects (and therefore less meaningful) are, the more versatile, customizable and able to be appropriated by a user they will be.

### 4.2 Experimental Protocol

The first step of the experimental protocol was to design the objects. It was made through participatory design sessions with musicians, designers and ordinary people. They were instructed to draw affordable objects for music applications described above (Fig. 1 provides a sample output). The second step was to pre-test these objects in order to validate their respective affordance. The results were validated by all participants.

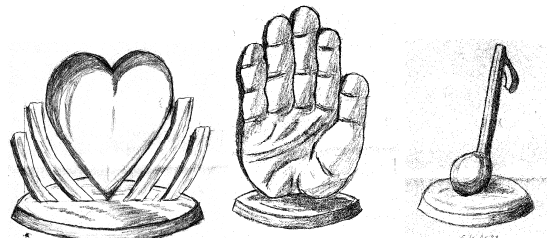


Figure 1. Object "rhythm" (heart shape), object "off" (hand-shaped) and object "note".

The interface objects were selected *a priori* from the "culture" induced by the graphical interfaces, like an eraser to erase, "stop" to stop "cursor" to the rhythm, etc. We will not describe in detail the design sessions that led to realize the matching items used in the remainder of the experiment. We took 16 subjects divided into 2 groups, each competing to test each hypothesis. For example, H1 task was to understand the function of objects in various situations. Some users were given less affordable objects while some others objects that provide good affordance properties. And so on for H'2 and H'3. Measures of time and number of errors were collected and completed by users interviews after the experiment using a Likert scale questionnaire

## 5. RESULTS

Results depending on the situation	Direct		Indirect	
	Obj. aff.	Obj. non aff.	Obj. aff.	Obj. non aff.
Understanding difficulty (Likert 1-6)	2, 0.5	3.5, 1.58	3.25, 1.6	3, 0.925
Understanding duration (in second)	10.5, 4.3	35.8, 24.2	36.3, 9.3	23.5, 11.4
Number of errors during the task	0.125, 0.3	0.375, 0.5	0.25, 0.4	0.6, 0.5

Figure 2. Results of measurements validated by questionnaires. The values in each cell are mean and standard deviation for all subjects who passed the test.

On this table (Fig.2), we can read that: in "direct situation" affordable objects are most relevant to the task (in terms of estimation of difficulty, time and numbers of errors): H1 is verified. Moreover, for an affordable object the "direct situation" is more favorable to the comprehension (in terms of errors, time and estimation of difficulty) and less favorable for the "indirect situation": H2 is verified. Everything happens as if there were an affordance "for itself" (egocentric) and affordance "for others" (exocentric). A closer analysis shows that other cues than the appearance can allow decoding the functions of non affordable objects (auditory cues, effects of their actions in the task, socio-cultural inferences). A corollary analysis also shows that H3 is validated. More precisely: the less affordable objects are and the more people can imagine it in another context (in terms of relevance for using the object in other context and estimation of the object affordance). Fig. 3 shows other aspects of affordance in relation to the type of objects (application object / interface object)

Results depending of the type of the objet	Application objet		Interface objet	
	Obj. aff.	Obj. non aff.	Obj. aff.	Obj. non aff.
Relevance for reusing the object in other context	2.625, 1.9	4, 2	6.75, 0.4	3.25, 1.4
Subjective estimation of the affordance ( Likert 1-6)	2.1, 0.8	6.5, 0.5	1.6, 0.3	5.5, 0.2

Figure 3. Results of measurements validated by questionnaires. The values in each cell are mean and standard deviation for all subjects who participated to the experiment.

## 6. DISCUSSION

The above results must be read with the usual precautions:

- a) The study is limited to a particular table and specific applications,

- b) The number of subjects of the experiment is limited and does not cover all socio-cultural categories,
- c) The conditions of the experiment are not entirely ecological.

Despite these restrictions, strong trends emerge due to the particularities of these objects "augmented": on the one hand the multi-functionality that appears in these objects is more difficult to grasp and understand for the subjects and, on the other hand, the question of "technological culture" raises other forms of affordance. Finally, the collaborative work seems to promote knowledge transfer and co-construction of meaning through some affordance we called exocentric.

## 7. CONCLUSION

The affordance of an object improves egocentric understanding of the task but it is not the only influence factor: the situation of collective work - that is to say what others do also with these objects - also plays a role in that phenomenon. Hence, the collective work improves the understanding of non affordable objects in the task. Moreover, affordance takes another and crucial signification within the framework of technological "augmented" objects for tangible interfaces: their multi-functional aspect makes these objects more difficult to understand but at the same time easier to transfer to other contexts by maintaining their functional property.

Socio-cultural factors are thus raised through affordance reuse and appropriation of objects.

## ACKNOWLEDGMENTS

This research was partially funded by the Ministry of National Education, Research and Technology, the CNRS, and especially the ANR (project TTT).

## REFERENCES

- Becker, M. (2010). *Le concept d'affordance et l'interface tangible Tangisense*. Mémoire de master, Université Paul Verlaine, Metz.
- Blackwell, A., Fitzmaurice, G., & Ishii, H. (2007). Tangible user interfaces in context and theory. *CHI 2007*, California, USA, May, ACM Press.
- Couture, N., Rivière, G., & Libération, D. (2007). Table interactive et interface tangible pour les géosciences : retour d'expérience. *Computer human interaction*, 3, 1-4.
- Fiebrink, R., & Morris, M. R. (2009). Dynamic Mapping of Physical Controls for Tabletop. *Computer Human Interaction*, 471-480.
- Fitzmaurice, G.W., Ishii, H., & Buxton, W. (1995). Bricks: laying the foundations for graspable user interfaces. *Proceedings of the SIGCHI conference on Human factors in computing systems*, Denver, Colorado, 442-449.
- Hornecker, E., & Buur, J. (2006). Getting a Grip on Tangible Interaction : A Framework on Physical Space and Social Interaction. *CHI 2006*, Québec, Canada, ACM Press.
- Ishii, H., & Ullmer, B. (1997). Tangible Bits: Towards Seamless Interfaces Between People, Bits, and Atoms. *CHI 1997*, 234-41.
- Kubicki, S., Lepreux, S., Kolski, C., & Caelen, J. (2010). Towards New Human-Machine Systems in contexts involving interactive table and tangible objects. *The 11th IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design, and Evaluation of Human-Machine Systems*, August 31 - September 3, Valenciennes, France.
- Kubicki, S., Lepreux, S., Lebrun, Y., Santos, P. D., Kolski, C., and Caelen, J. (2009) *New human-computer interactions using tangible objects: Application on a digital tabletop with rfid technology*. In Jacko, J. A. (Ed.), *Human-Computer Interaction*, volume 5612 of *Lecture Notes in Computer Science*, pages 446-455. Berlin: Springer.
- Moggridge, B. (2006). *Designing interactions*. MIT university Press.
- Mucchielli, A. (2000). *L'art d'influencer*. Paris : Ed. Armand Colin.
- Norman, D. A. (1999). Affordances, conventions and design, *Interactions*. *ACM*, 38-42

# SHOW-AND-TELL PLAY-IN: COMBINING NATURAL LANGUAGE WITH USER INTERACTION FOR SPECIFYING BEHAVIOR

Michal Gordon and David Harel  
*Weizmann Institute of Science*  
*Rehovot, Israel*

## ABSTRACT

In search of improving the ways to create meaningful systems from requirements specifications, this paper combines the *showing* and *telling* of how a system should behave. Using scenario-based programming and the language of *live sequence charts*, we suggest how user interaction with the system and user written requirements in natural language can interleave to create specifications through an interface that is both natural and agile.

## KEYWORDS

Intelligent interfaces, Requirement engineering, Scenario-based programming, Live sequence charts

## 1. INTRODUCTION

Scenario-based programming is a method that allows specifying system behavior by describing system scenarios using precise and executable methods. The language of *live sequence charts* (LSC) (Damm and Harel 2001) is one method for these types of descriptions. LSCs add expressive power to earlier sequence-based languages by being multi-modal: an LSC can distinguish what must happen from what may happen, and can specify also what is forbidden from happening. The resulting specification is fully executable.

One of the advantages of LSCs is their use for describing system behavior for reactive systems. The language constitutes a step in the direction of *liberating programming* and making programming more accessible to people who are not programmers, as described in (Harel 2008). The LSC language has been extended with a tool (the *Play-Engine*) that supports intuitive GUI-based methods for capturing the behavior (termed *play-in*) and for executing a set of LSCs (termed *play-out*); see (Harel and Marelly 2003). The present work focuses on introducing an enriched method for play-in, which creates an improved interface for specifying system requirements and for scenario-based programming.

The new method combines natural language parsing methods with user interaction and uses these to create an intelligent user interface. The user specifying the system's behavior can use the method most relevant for the type of behavior he/she is specifying, by *showing* — interacting with the system or by *telling* — describing (parts of) the scenario in a semi-natural language (Gordon and Harel 2009). Any textual requirements thus entered are parsed, so that our *show-and-tell* (*S&T*) play-in algorithm can intelligently guess the user's intention when there are multiple possibilities.

As in real life, a picture is often worth a thousand words and other times a textual description is more appropriate. In analogy, there may be cases when the interaction is simpler to put in words than to demonstrate, or vice-versa. The main contribution of this paper is in combining the two in a natural and semantically meaningful way.

## 2. THE LANGUAGE OF LSCS AND PLAY-IN

In its basic form, an LSC specifies a multi-modal piece of behavior as a sequence of message interactions between object instances. It can assert mandatory behavior — what must happen (with a *hot* temperature) —

or possible behavior — what may happen (with a *cold* temperature), as well as what is forbidden from happening. In the LSC language, objects are represented by vertical lines, or lifelines, and messages between objects are represented by horizontal arrows between objects. Time advances along the vertical axis and the messages entail an obvious partial ordering. The events that trigger the scenario appear at the top in blue dashed lines; if they are satisfied, i.e., all events occur and in the right order, then the hot events (in red solid lines) must be satisfied too. See (Harel&Marely 2003).

Play-in is a method for capturing a scenario in an LSC by interacting with a GUI representation of the system. This allows users to operate the final system, or a representation of it, thus 'recording' their behavior by demonstrating it. Although play-in is intuitive and can be easily adopted by users without orientation to programming, it has some drawbacks. First, it requires a pre-prepared GUI of the non-behaving system. Although this is reasonable for some systems (e.g., a general robot with no behavior), in others the specification of the GUI will typically only emerge after considering parts of the system behavior. Another problem is that interactions relevant to the system's behavior often take place among logical objects, for which there can exist only some general representation with possible lower level functions and properties. It is straightforward to interact with a button that initializes a wireless connection, but there is no need to force a graphical representation on the wireless connection when referring to it.

There are also many requirements that are less interactive and more programmatic in their essence, such as specifying a condition or selecting some variable. These are the cases where S&T-play-in becomes relevant: natural language descriptions are quick and simple and can be interleaved with interaction when it is most relevant, as we show below.

Play-in has been extended by controlled natural English (Gordon and Harel 2009). Nouns and verbs are found using the Wordnet dictionary by Miller et. al. (1993) and LSCs are created based on interactions specified in clear sentences, using a specially tailored context free grammar for LSCs. Ambiguities are resolved by the person entering the requirements when the sentences do not translate into meaningful LSCs. The model of the system, its objects and possible low level behaviors, accumulate, and are used to translate further sentences more successfully. In this scheme, no GUI is required at the initial stages of the process, and it can be designed later, based on the system model. This allows the requirements engineer or programmer to concentrate on the system's behavior rather than on its structure and components.

One advantage of the natural language approach is the ability to refer to system objects, conditions, variables and loops in the text in a way that is close to the process performed when the application expert simply writes what he/she wants the system to do.

The natural language play-in of Gordon and Harel (2009) emphasizes writing logical constructs in English, rather than selecting them from menus or dragging them from a graphical toolbar. However, there are cases when using the mouse to point and select is quicker. When a certain knob has a graphical representation and possible low level behaviors, then *showing* the action may be more convenient than *telling* or describing it textually. As when a parent directs a child to return the milk to its proper place in the refrigerator could involve the parent saying 'please return the milk to its place', while pointing to the refrigerator.

### 3. SHOW-AND-TELL PLAY-IN

The *show-and-tell-play-in* method (*S&T-play-in*) uses online parsing and the state of the current parse to interpret the interaction and integrate it into the scenario-based requirement; i.e., into the LSC that is being constructed on the fly.

An interaction can be interpreted in multiple ways. When an object is selected (from the model or the GUI), either its name or the operation performed on it (e.g., clicked, or turned) may be integrated into the sentence. When an object property is selected, it may be a reference to the property name or to the property value. The parsing is performed bottom-up using an active chart parser similar to that of Kay (1986) and adapted for online parsing as in Jurafsky and Martin (2009), Figure 1a shows the system architecture and Figure 2 provides details of the algorithm.

In each requirement being entered, the indexes represent the locations between the words (as in  $_0$  when  $_1$  the  $_2$  user  $_3$ ). An edge represents a grammar rule and the progress made in recognizing it. We use the common dotted rule, where a dot ('•') within the right-hand side of the edge indicates the progress made in recognizing

the rule, and two numbers indicate where the edge begins on the input and where its dot lies.

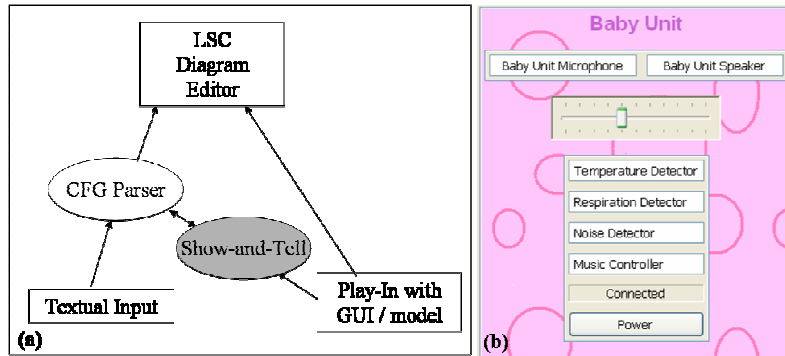


Figure 1. (a) The system architecture. (b) Part of the baby monitor sample application GUI.

Technically, an interaction generates possible edges for the parsing with the object names or operations selected, and the algorithm tests which edges complete the current parse properly. The longest valid completion is selected (Figure 2b), and additional valid possibilities are presented to the user and can be selected by him/her on the fly.

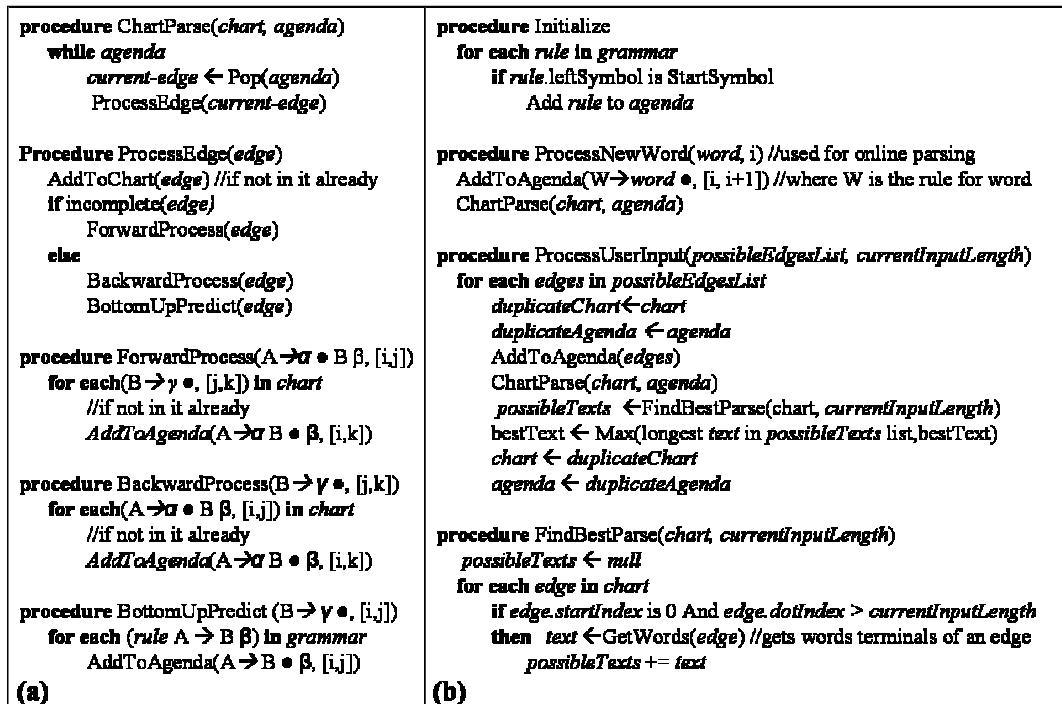


Figure 2. (a) Parse procedures from Jurafsky and Martin (2009), (b) *ProcessNewWord* is used for online parsing, while *ProcessUserInput* is used to fuse interactions into the parsing.

Consider the example in Fig. 3. The textual requirement at the interaction point is ‘when the user’, so the parse is not complete. However, some edges are already completed in the bottom-up parsing, shown above the sentence part, as can be seen in the left part of the figure, which displays edges as curved lines.

We assume that an interaction creates only complete edges, since when guessing what the user meant, it is reasonable to assume he/she thinks in complete ‘chunks’ of the language; e.g., he/she can refer to a noun, a verb, or parts of sentences that include them. For example, while ‘[the button]’ and ‘[clicks] [the button]’ are reasonable edges and are complete edges in the grammar, ‘[clicks] [the]’ is not, as can be seen in Fig. 3c.

At each step, the algorithm adds one of the edges or a set of edges to the current parse, and tests whether these interaction edges advance or complete any of the parse edges, as shown in Fig. 2 *ProcessUserInput*.

From the possible completed or advanced edges, the longest one is selected; in Fig. 3b, this would be the top edge.

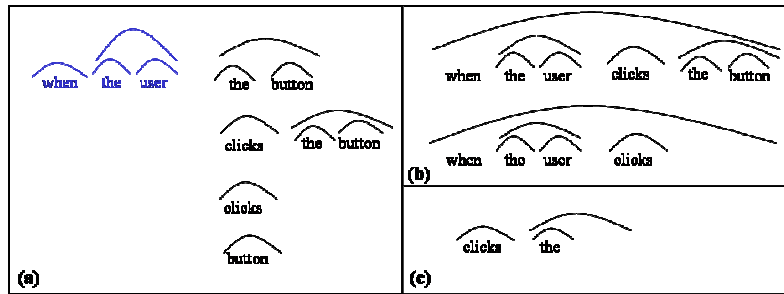


Figure 3. Parse sample for the sentence "when the user" and an interaction of [clicking a button].

#### 4. CASE STUDY: THE BABY MONITOR

We describe some examples from the development process of a baby monitor system, which allows parents to watch over their baby by monitoring respiratory movement and room temperature. We depict interaction outputs in square brackets.

Since the parsing is online and the transformation to LSCs is linear, complete constructs can be directly transformed to their LSC counterparts, which could allow the user to see the LSC created as he/she works. Our current implementation only adds text according to user interaction and generates the visual representation of the LSC when the user completes the requirement and selects the generate LSC option.

The interaction with a GUI (Figure 1b) or with the system model (the list of system objects without their graphical representations) can involve one of three actions: selecting the object, performing an action (e.g., calling a method) or setting/getting a property (attribute) of the object.

One issue that needs to be dealt with when specifying LSC requirements is the question of whether the interaction is meant as a full interaction or just the selection of the object. In the sentence: "*when the user clicks the [increase-threshold-temperature-button], the temperature-threshold increases*", the interaction of clicking the button adds only the button name to the already entered text. However, another example that adds a full phrase is: "*when [the user drags the sensitivity-button], the sound sensitivity changes to the sensitivity-button value*". Notice that the interaction adds a full phrase of dragging the sensitivity-button and not only the object name because of the different text entered when the interaction occurs.

Another type of information that can be entered is the selection of properties, as in the sentence: "*when the baby-unit temperature changes, if the baby-unit temperature is greater than [temperature-threshold], the [light state changes to blinking]*". In the first interaction, only the threshold property itself is added by the interaction, while in the second part, the user changes the state of the light to blinking and the full phrase is added to the sentence. Sample clips can be found in <http://www.wisdom.weizmann.ac.il/~michalk/SaT/>.

#### 5. RELATED WORK

The work presented here builds upon the original play-in idea of (Harel and Marelly, 2003), which allows user interaction with a GUI for specifying behavior. User interaction for capturing behavior is also found in many programming-by-example systems, from programming by dragging icons on screen in the Pygmalion, Cocoa or Stagecase environments to constructing grammars with the Grammex system, all described by Cyper et.al. (1993). These systems can be viewed as extensions of macro systems that allow recording a sequence of operations performed by the user and then repeating the sequence while generalizing some aspects of the operations, rather than specifying the full behavior explicitly.

The idea of multimodal interfaces as discussed by Ingebreetsen (2010) is already in use in intelligent interfaces for gaming and in smart phones. These modalities include speech, facial expression, body posture, gestures and bio-signals. The question of multimodal synchronization and fusion is interesting for many



application areas. Perhaps the initial methods we discuss for requirement engineering synchronizing text (or speech) and user's mouse operations (comparable to gestures) can be useful in other fields too.

Using speech recognition and natural language as an interface for specification has been discussed before. For example, in (Graefe and Bischoff, 1997), interacting with a robot can benefit from a combination, where the context, the knowledge base acquired by the robot at each point, helps to better understand the directions to the robot. Our method currently parses textual requirements, but we have also tested the use of speech recognition dictation with the Microsoft™ Speech API 5.1. In S&T, the text (or speech) is used as the context for understanding the interaction.

## 6. CONCLUSION AND FUTURE WORK

This paper introduces the idea of combining two interfaces: text and interaction with a GUI into a single intelligent interface that interprets an interaction based on the state of the textual parse to allow generating system requirements in a natural way. We show here only some of the possibilities of combining interaction and text, as the domain of possibilities is fixed by the system interface and the grammar. Other interfaces that have a text/speech interface using a grammar or command-and-control may benefit from a suitably adapted version of the S&T interface.

The method requires further evaluation. One way to do this is to check and compare the time and effort required to create diagrams using only play-in, using only natural language play-in and using S&T-play-in.

One extension we would like to add to the algorithm, is a generic method to create interaction edges using a given grammar and minimal information on the interaction. This may make the method more useful outside the particular LSC-based method for specifying system behavior.

## ACKNOWLEDGEMENTS

We would like to thank Smadar Szekely and the other members of our group's software team developing the PlayGo tool used to test the method. This research was supported in part by the John von Neumann Minerva Center for the Development of Reactive Systems at the Weizmann Institute of Science, and by an Advanced Research Grant from the European Research Council (ERC) under the European Community's Seventh Framework Programme (FP7/2007-2013).

## REFERENCES

- Cypher, A. et al, 1993. *Watch What I Do: Programming by Demonstration*. MIT Press.
- Damm, W. and Harel, D., 2001. LSCs: Breathing Life into Message Sequence Charts. *Formal Methods in System Design* 19:1, 45–80.
- Gordon, M. and Harel, D. 2009. Generating Executable Scenarios from Natural Language. *Proc. 10th Int. Conf. on Computational Linguistics and Intelligent Text Processing (CICLing'09)*, Mexico, pp. 456–467.
- Graefe, V. and Bischoff, R., 1997. A Human Interface for an Intelligent Mobile Robot. *Proc. 6<sup>th</sup> Int. Workshop on Robot and Human Communication*. Japan, pp. 194–199.
- Harel, D. 2008. Can Programming be Liberated, Period?, *IEEE Computer* 41:1, 28–37
- Harel, D. and Marelly R., 2003. Specifying and Executing Behavioral Requirements: The Play In/Play-Out Approach. *Software and System Modeling* 2, 82–107.
- Ingebretsen, M. 2010, In the News, *Intelligent Systems, IEEE* 25:4, 4–8
- Jurafsky, D. and Martin, J. H. 2009, *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition*, Pearson Prentice Hall.
- Kay, M., 1986. Algorithm schemata and data structures in syntactic processing. In *Readings in Natural Language Processing*, Morgan Kaufmann, pp. 35–70.
- Miller, G.A., Beckwith, R., Fellbaum, C., Gross, D. and Miller, K., 1993. Introduction to WordNet: An On-line Lexical Database. <http://wordnet.princeton.edu/>.

# EVALUATING EFFECTS OF VISUAL APPEARANCE OF VIRTUAL HUMANS IN E-COMMERCE

Rainer Blum<sup>1</sup> and Malcolm Clarke<sup>2</sup>

<sup>1</sup>*Hochschule Fulda - University of Applied Sciences, Dept. of Applied Computer Science  
Marquardstr. 35, 36039 Fulda, Germany*

<sup>2</sup>*Brunel University, Dept. of Information Systems and Computing  
Uxbridge, Middlesex, UB8 3PH, UK*

## ABSTRACT

Customized avatars as imitations of individual, real customers constitute a promising concept for personalized product marketing in Internet retail. But, consolidated knowledge about effective visual design strategies of these surrogates to positively influence consumer behavior is still lacking. This work-in-progress paper presents and discusses an experimental study targeted at this gap. It focuses especially on the composition of suitable outcome measures, covering central consumer behavior-related reactions and attitudes and takes the experiential versus utilitarian characteristics of this marketing concept into account.

## KEYWORDS

e-commerce, avatar design, consumer behavior, user requirements, experiment design

## 1. INTRODUCTION

In recent years B2C e-commerce settings have experienced an increasing use of graphical representation of humans as one of the many additional, often interactive, website features. The predominant purpose or role of these virtual humans is to act as virtual sales agents. At the same time, the value of these customized virtual humans in the form of digital representations, or embodiments of individual, real persons, so-called avatars, is an important aspect of current e-commerce research. Furthermore, such 'virtual twins' of consumers are expected to gain an increasing importance in the future – as appearance and behavior of avatars becomes ever closer to their human counterpart. Individualized avatars constitute the focus of this research. Here, the avatars provide visualization and information of the product by using the individual appearance of the user. A typical use is the virtual try-on of clothing, in which customer-individualized virtual humans wear garment products. This virtual try-on application is the subject of this research.

According to Pena et al. (2009), digital self-representations of people affect their cognition and behavior, making it possible to increase the probability that they think and behave in predictable ways without arousing suspicion. The issue is to determine which are the most important characteristics of an avatar that is intended for consumer-centered, customized product representation in a retail shopping context. Moreover requirements must take into account that the self-evident, ultimate purpose of such a marketing measure is to support product sales. Although this goal originates from the viewpoint of the retailers, the avatar design must be geared to the effects on the attitudes and behavior of the consumers.

This research focuses exclusively on the visual aspects of the presentation, rather than factors such as the imitation of individual behavior of the virtual human. This approach has been adopted in the majority of the existing work on the topic: Calhoun et al. (2007) demonstrate "My Virtual Model" (MVM), which is a popular virtual try-on solution utilizing personalized avatars, to be able to raise online turnover, because consumer confidence in purchasing is improved. They also presented participants with their high-resolution body scans, displayed as a 3D model but without color and texture information. Participants voiced many criticisms of the MVM avatars, particularly that the body was too smooth, lacked variation in appearance with respect to the input of body measurements and that it looked more like a model than a real person, resulting in an unrealistic image. In contrast, many participants stated that their body scan was too detailed

and realistic. Furthermore the study of Calhoun et al. (2007) was not conducted using a systematic variation of avatar appearance for comparison.

Merle et al. (2009) tested several degrees of personalization within the options offered by the MVM technology. The experimental condition with the highest level of personalization involved consumer face imitation. The MVM face feature was a simple overlay of the person's face cut from an uploaded 2D photo onto the avatar, making the individual face only available in the front view. However participants showed highest values for the investigated consumer behavior-related constructs for this most personalized model. Consequently, the authors recommend maximization of personalization.

In contrast, in a study on similarity and self-identity in advertising by Lutchyn et al. (2009), participants rated models (represented only as the face) that were less similar in appearance to them (concerning ethnic matching) as more trustworthy, attractive, likeable and even as more similar to themselves. As explanation for this effect, the authors suggest, referring to an assimilation-contrast reaction according to Mussweiler's (2003) selective accessibility model, that the degree of similarity acts as a cue whether to compare or contrast oneself to the model. In the case of high congruence, people compare themselves and therefore, mainly consider differences and thus the inappropriateness of the model. When differences are dominant the similarities come to the fore resulting in a more favorable attitude to the model.

In summary, existing research on the requirements for personalized virtual humans – taking the role of digital consumer representations – and the resulting effects on consumer behavior is limited and somewhat inconsistent. At least, it seems evident that different visual characteristics have different effects on consumer behavior. However, more systematic and comprehensive studies to clarify how to achieve desired effects and to consolidate existing findings are advisable. For example, an explicitly abstract appearance condition has not so far been considered in existing research. Consequently, the main research question of this work is how attributes of visual avatars affect consumer behavior, to what extent this happens and which attributes most contribute to these effects. It is hypothesized that visual characteristics of avatars can be used systematically to significantly influence consumer behavior.

This paper presents and discusses the approaches for a user study to elicit factors and determine their relative importance. The intent is to broaden understanding of cause-effect relations in avatar design for e-commerce and to provide new and greater insight on the properties of this prospective concept for (personalized) product marketing.

## 2. EXPERIMENTAL DESIGN

A prototype, but typical B2C-style online shop is used as stimulus. It offers contemporary and fashionable women's clothing items. The usual product catalogue and shopping cart is augmented with a 'virtual try-on' feature. It provides female participants a highly realistic try-on of clothing items in a 3D virtual reality space by dressing a personalized avatar with digitized versions of the garments.

Each participant is randomized to one of three groups that vary the characteristics of the avatar: (i) high personalization (individualized head and face imitation, photorealistic style); (ii) typical, non-famous fashion model (idealized, photorealistic style); and (iii) an abstract look-and-feel of the avatar. These conditions cover a broad spectrum of fundamental categories of visual appearance. The core characteristics of the appearance of each avatar type are elaborated clearly and with comparable quality. The study utilizes unique, high-quality visualizations in true 3D, so far not present in similar research. Daugherty et al. (2008) demonstrated the advantage of 3D product presentation over traditional forms of product advertising.

The participants in each group are asked to complete a typical shopping task with multiple parts using the stimulus shop, and with a focus on the use of the try-on of clothing using the avatar. The aim is to allow the participants to experience the avatar concept, to familiarize them with it and give them the opportunity to develop an individual attitude towards its use, asking them to take the role of potential customers. The aim is to elicit the most realistic reactions possible in this simulation.

Although the avatars have different design characteristics, they retain the individual body measurements of the customer. This ensures a comparable 3D presentation of the clothing across the groups, so that the three conditions differ only in the characteristics of the avatars.

After experiencing the shop and their personal avatar, the participants are asked to complete a questionnaire comprising a number of psychometric, Likert-style scales, designed to measure specific consumer behavior-related, attitudinal constructs. This is followed by an interview.

### 3. DATA COLLECTION

Our experiment to determine and investigate factors that positively affect purchase concentrates on purchase-related consumer behavior. The literature describes many relevant constructs about reactions and attitudes towards marketing stimuli. We have therefore selected well-investigated effects on purchase behavior, ranked them, and selected the most significant as the variables for the theoretical model in this research. We have also added experiential factors (entertaining, playful) in addition to the utilitarian character and qualities of the personalized avatar try-on feature.

The first category is attitudinal measures. A psychometric scale is used for each of the three attitudes differentiated in the common ABC model of attitudes: 'Shopping enjoyment' as affective response; 'purchase intention' as behavioral intention; and 'decision support satisfaction' as cognitive response; plus 'attitude towards the website' as a general attitudinal evaluation.

'Shopping enjoyment' constitutes a person's enjoyment of the shopping experience. The corresponding scale measures the extent to which a person perceives the shopping experience with an online shop to be enjoyable in its own right, apart from any performance consequences that may be expected. 'Purchase intention' represents what and where a person thinks she or he will buy. This construct is generally considered to be highly indicative for actual purchase behavior. 'Decision support satisfaction' is chosen as cognitive response as it evaluates how satisfied a person is with the decision support received, including the ability of the system to deliver relevant information for decision making and to improve the person's decisions. Finally, attitude towards the website represents a person's general evaluation of a website or general favorability towards a website. Note that the items of those scales that refer to "website" are adapted to target towards the "try-on with personal avatar" scenario studied here.

The second category is the utilitarian value of personalized avatar augmented clothing shopping and is gauged via the 'website informativeness' construct. The respective scale measures the degree to which a website provides a person with information she or he perceives as resourceful and helpful. Furthermore, as the purpose of the personalized avatar is to emulate the product experience that is possible and common in a retail store, the 'telepresence' construct is taken into account. This provides a subjective perception of how closely the online sensory information and interaction with the product (with the help of and via the personalized avatar) approximates information and interaction with the real product in a brick-and-mortar store.

The third category of purchase-related consumer behavior is 'perceived risk', and relates to the nature and amount of uncertainty perceived by a person in making a particular purchase decision. It constitutes a central issue in Internet retail and is known to influence highly the intention to shop online: As product perception by means of physical inspection is not possible, online consumers suffer from evaluation difficulty – a cognitive and behavioral difficulty and effort required to judge and discriminate among alternatives. Good information content is essential for perceived risk reduction.

A 'self-congruence' scale is employed to operationalize the congruence between avatar and own self-concept. For this experiment, it is assumed that this 'avatar similarity to self' dimension varies consistently between groups. This might also provide an explanation for differences in outcome between the investigated conditions as well as hints for avatar design. However, differences in attitude towards the personalized avatar may also be due to other factors, such as age or habits and experience concerning use of Internet and online shopping (especially shopping of clothing). This data is collected for further statistical analysis.

Finally, a construct called 'shopping orientation' is used that represents the manner in which a person approaches the activity of shopping. The chosen scale produces a value on a continuum between recreational shopping orientation and task orientation, and corresponds to the experiential versus utilitarian character of the personalized avatar feature. For the same purpose, an 'appearance orientation' scale is used. It gauges the importance of appearance to the individual, with attention to appearance as well as thought and behavior centered on appearance. A possible effect on the above consumer behavior factors as well as the 'self-congruence' construct shall be analyzed.

In addition to the questionnaire, all participants are interviewed with open-ended questions. Purpose is to explore the motivations for the responses given in the questionnaire and to seek explanations for identified effects or trends. The main aspects of interest are the like and dislike of different attributes and aspects of the avatar, preferred characteristics of the avatar and to assess the perceived purpose and influence of the avatar (experiential vs. utilitarian value, decision to purchase).

#### 4. CONCLUSION

The presented procedure is designed as a comprehensive approach, exploring central aspects that may guide avatar design in the specific e-commerce context. The experiment design implements a systematic variation of avatar appearance, varying from high to no (or minimal) personalization and photorealism. The three conditions cover and contrast a broad spectrum of fundamental visual appearance categories. The specific measures were selected and compiled in such a way as to include the factors determined most decisive and representative of the relevant spectrum of consumer behavior aspects. Therefore, valid and significant findings may be expected. The experimental procedure was pre-tested for effectiveness with several persons.

Limitations could arise from the artificial nature of the experimental scenario. The participants take the role of potential customers in order to elicit opinion, but they are not in a real shopping situation. Due to resource restrictions only a relatively small number of participants can be included (20 per group), although this number is anticipated to yield statistically significant results from its power calculation. Furthermore, although care has been taken over the design of the experiment and choice of participant, other factors outside the characteristics of the avatar may vary and affect results.

The results will inform practical design and guidelines shall be derived in order to facilitate central, practical decisions when developing such avatar applications. They can provide guidance for the question, which are the most important visual characteristics for an avatar that is intended for consumer-centered, customized product presentation, and cover aspects including: the degree of photorealism; characteristics of the consumers to be imitated; level of simplification of detail; and look-and-feel. These findings can also help to identify technical aspects of the generation of avatars and where future development effort is worthwhile. The latter is significant, as today, high-quality, customized avatars are not yet ready for use in the e-commerce domain due to missing automation and insufficient integration of available, even if partial, solutions.

#### ACKNOWLEDGEMENT

This research was financially supported by the German Federal Ministry of Education and Research within the framework of the program “Forschung an Fachhochschulen mit Unternehmen (FHprofUnd)” (project acronym “KAvaCo”, no. 1702X09).

#### REFERENCES

- Calhoun, E., et al., 2010. Virtual fit of apparel on the internet: Current technology and future needs. *Handbook of research in mass customization and personalization*. Eds. F. T. Piller & M. M. Tseng. World Scientific, pp. 731–748.
- Daugherty, T., et al., 2008. Consumer learning and the effects of virtual experience relative to indirect and direct product experience., *Psychology and Marketing*, vol. 25, no. 7, pp. 568–586.
- Lutchyn, Y., et al., 2009. Ethnic Matching: An Examination of Ethnic Morphing In Advertising. Virtual social identity and consumer behavior. Eds. N. T. Wood & M. R. Solomon. M.E. Sharpe, Armonk, N.Y.
- Merle, A., et al., 2009. Do I Recognize Myself in this Avatar? An Exploratory Study of Self-Congruity and Virtual Model Personalization Level. *5th World Conference on Mass Customization and Personalization. Helsinki, Finland*.
- Mussweiler, T., 2003. Comparison Processes in Social Judgment: Mechanisms and Consequences., *Psychological Review*, vol. 110, no. 3, pp. 472–489.
- Pena, J., et al., 2009. The Priming Effects of Avatars in Virtual Settings. *Communication Research*, vol. 36, no. 6, pp. 838–856.

# A FACIAL ANIMATION FRAMEWORK WITH EMOTIVE/EXPRESSIVE CAPABILITIES

G. Riccardo Leone and Piero Cosi

*Institute of Cognitive Sciences and Technologies – National Research Council  
Via Martiri della libertà 2, 35137 Padova, Italy*

## ABSTRACT

LUCIA is an MPEG-4 facial animation system developed at ISTC-CNR. It works on standard Facial Animation Parameters and speaks with the Italian version of FESTIVAL TTS. To achieve an emotive/expressive talking head LUCIA was built from real human data physically extracted by ELITE optotracking movement analyzer. LUCIA can copy a real human by reproducing the movements of passive markers positioned on his face and recorded by the ELITE device or can be driven by an emotional XML tagged input text, thus realizing a true audio/visual emotive/expressive synthesis. Synchronization between visual and audio data is very important in order to create the correct WAV and FAP files needed for the animation. LUCIA's voice is based on the ISTC Italian version of FESTIVAL-MBROLA packages, modified by means of an appropriate APML/VSML tagged language. LUCIA is available in two different versions: an open source framework and the "work in progress" WebGL

## KEYWORDS

Talking head, facial animation, mpeg4, affective computing, TTS, WebGL

## 1. INTRODUCTION

There are many ways to control a synthetic talking face. Among them, geometric parameterization [Massaro et al., 2000], morphing between target speech shapes [Bregler et al., 1997], muscle and pseudo-muscle models [Terzopoulos et al., 1995 - Vatikiotis-Bateson et al., 1996], appear the most attractive.

Growing interest have encountered text to audiovisual systems [Beskow, J., 1995 - LeGoff, B. and Benoit, C., 1996], in which acoustical signal is generated by a Text to Speech engine and the phoneme information extracted from input text is used to define the articulatory movements.

To generate realistic facial animation is necessary to reproduce the contextual variability due to the reciprocal influence of articulatory movements for the production of following phonemes. This phenomenon, defined co-articulation [Farnetani and Recasens, 1999], is extremely complex and difficult to model. A variety of coarticulation strategies are possible and even different strategies may be needed for different languages [Bladon and Al-Bamerni, 1976]. A modified version of the Cohen-Massaro co-articulation model [Cosi and Perin, 2002] has been adopted for LUCIA [Fusaro et al., 2003] and a semi-automatic minimization technique, working on real cinematic data acquired by the ELITE opto-electronic system [Ferrigno and Pedotti, 1985], was used for training the dynamic characteristics of the model, in order to be more accurate in reproducing the true human lip movements.

Moreover, emotions are quite important in human interpersonal relations and individual development. Linguistic, paralinguistic and emotional transmission are inherently multimodal, and different types of information in the acoustic channel integrate with information from various other channels facilitating communicative processes. The transmission of emotions in speech communication is a topic that has recently received considerable attention, and automatic speech recognition (ASR) and multimodal or audio-visual (AV) speech synthesis are examples of fields, in which the processing of emotions can have a great impact and can improve the effectiveness and of human-machine interaction.

Viewing the face improves significantly the intelligibility of both natural and synthetic speech, especially under degraded acoustic conditions. Facial expressions signal emotions, add emphasis to the speech and

facilitate the interaction in a dialogue situation. From these considerations, it is evident that, in order to create more natural talking heads, it is essential that their capability comprises the emotional behavior.

In our TTS (text-to-speech) framework, AV speech synthesis, that is the automatic generation of voice and facial animation from arbitrary text, is based on parametric descriptions of both the acoustic and visual speech modalities. The visual speech synthesis uses 3D polygon models, that are parametrically articulated and deformed, while the acoustic speech synthesis uses an Italian version of the FESTIVAL diphone TTS synthesizer [Tesser, F. et al., 2001] now modified with emotive/expressive capabilities. The block diagram of our framework is depicted in Fig. 1a.

Various applications can be conceived by the use of animated characters, spanning from research on human communication and perception, via tools for the hearing impaired, to spoken and multimodal agent-based user interfaces. The recent introduction of WebGL, which is 3D graphics in web browsers, opens the possibility to bring all these applications via internet. A software porting of LUCIA facial animation framework is currently in development.



Figure 1. a) LUCIA's functional block diagram, b) position of reflecting markers and reference planes for the articulatory movement data collection (*on the left*), and the MPEG-4 standard facial reference points (*on the right*)

## 2. DATA ACQUISITION ENVIRONMENT

LUCIA is totally based on true real human data collected during the last decade by the use of ELITE [Cosi and Magno-Caldognetto, 1996 - Zmarich et al., 1997 - Magno-Caldognetto et al, 1998], a fully automatic movement analyzer for 3D kinematics data acquisition, which provides for 3D coordinate reconstruction, starting from 2D perspective projections, by means of a stereo-photogrammetric procedure which allows a free positioning of the TV cameras. The 3D data dynamic coordinates of passive markers (see Fig. 1b) are then used to create our lips articulatory model and to drive directly our talking face, copying human facial movements.

Two different configurations have been adopted for articulatory data collection: the first one, specifically designed for the analysis of labial movements, considers a simple scheme with only 8 reflecting markers (bigger markers in Fig.1b) while the second, adapted to the analysis of expressive and emotive speech, utilizes the full and complete set of 28 markers. All the movements of the 8 or 28 markers, depending on the adopted acquisition pattern, are recorded and collected, together with their velocity and acceleration, simultaneously with the co-produced speech which is usually segmented and analyzed by means of PRAAT [Boersma, 1996], that computes also intensity, du-ration, spectrograms, formants, pitch synchronous F0, and various voice quality parameters in the case of emotive and expressive speech [Magno Caldognetto et al., 2003 - Drioli et al., 2003].

In order to simplify and automates many of the operation needed for building-up the 3D avatar from the motion-captured data at ISTC-CNR we developed INTERFACE [Tisato et al, 2005], an integrated software designed and implemented in Matlab©

## 3. SYSTEM ARCHITECTURE

LUCIA is a graphic MPEG-4 standard compatible facial animation engine implementing a decoder compatible with the "Predictable Facial Animation Object Profile". LUCIA speaks with the Italian version of FESTIVAL TTS, as illustrated in Fig. 1a.

MPEG4 specifies a set of Face Animation Parameters (FAPs), each corresponding to a particular facial action deforming a face model in its neutral state. A particular facial action sequence is generated by deforming the face model, in its neutral state, according to the specified FAP values, indicating the magnitude of the corresponding action, for the corresponding time instant. Then the model is rendered onto the screen.

LUCIA is able to generate a 3D mesh polygonal model by directly importing its structure from a VRML file [Hartman and Wernecke, 1996] and to build its animation in real time.

Currently the model is divided in two sub sets of fundamental polygons: the skin on one hand and the inner articulators, such as the tongue and the teeth, or the facial elements such as the eyes and the hair, on the other. This subdivision is quite useful when animation is running, because only the reticule of polygons corresponding to the skin is directly driven by the pseudo-muscles and it constitutes a continuous and unitary element, while the other anatomical components move themselves independently and in a rigid way, following translations and rotations (for example the eyes rotate around their center). According to this strategy the polygons are distributed in such a way that the resulting visual effect is quite smooth with no rigid "jumps" over all the 3D model.

LUCIA emulates the functionalities of the mimic muscles, by the use of specific "displacement functions" and of their following action on the skin of the face. The activation of such functions is determined by specific parameters that encode small muscular actions acting on the face, and these actions can be modified in time in order to generate the wished animation. Such parameters, in MPEG-4, take the name of Facial Animation Parameters and their role is fundamental for achieving a natural movement. The muscular action is made explicit by means of the deformation of a polygonal reticule built around some particular key points called "Facial Definition Parameters" (FDP) that correspond to the junction on the skin of the mimic muscles.

Moving only the FDPs is not sufficient to smoothly move the whole 3D model, thus, each "feature point" is related to a particular "influence zone" constituted by an ellipses that represents a zone of the reticule where the movement of the vertexes is strictly connected. Finally, after having established the relationship for the whole set of FDPs and the whole set of vertexes, all the points of the 3D model can be simultaneously moved with a graded strength following a raised-cosine function rule associated to each FDP.

There are two current versions of LUCIA: an open source 3D facial animation framework written in C programming language (see <http://sourceforge.net/projects/lucia/>) and a new WebGL implementation. The C framework allows efficient rendering of a 3D face model in OpenGL-enabled systems (it has been tested on Windows and Linux using several architectures), has a modular design and provides several common facilities needed to create a real-time Facial Animation application. The very recent introduction of 3D graphics in the web browsers (which is known as WebGL) opens new possibilities for our 3D avatar. The powerful of this new technology is that you don't need to download any additional software or driver to access the content of the 3D world you are interacting with. We are currently developing this new software version in order to easily integrate LUCIA in a website and use her as a virtual guide for the wikimemo.it project - The portal of Italian Language and Culture.

## 4. EMOTIONAL SYNTHESIS

Audio Visual emotional rendering was developed working on true real emotional audio and visual databases whose content was used to automatically train emotion specific intonation and voice quality models to be included in FESTIVAL Italian TTS system [Tesser et al., 2004 - Tesser et al, 2005 - Drioli et al., 2005.- Nicolao et al., 2006] and also to define specific emotional visual rendering to be implemented in LUCIA [Cosi et al., 2004 - Magno-Caldognetto et al., 2004 - Cavicchio et al, 2004].

An emotion specific XML editor explicitly designed for emotional tagged text was developed. The APMML mark up language [Pelachaud et al., 2004] for behavior specification permits to specify how to markup the verbal part of a dialog move so as to add to it the "meanings" that the graphical and the speech generation components of an animated agent need to produce the required expressions.

So far, the language defines the components that may be useful to drive a face animation through the facial description language (FAP) and facial display functions. The extension of such language is intended to support voice specific controls. An extended version of the APMML language has been included in the FESTIVAL speech synthesis environment, allowing the automatic generation of the extended phonation file



from an APMML tagged text with emotive tags. This module implements a three-level hierarchy in which the affective high level attributes (e.g. <anger>, <joy>, <fear>) are described in terms of medium-level voice quality attributes defining the phonation type (e.g., <modal>, <soft>, <pressed>, <breathy>, <whispery>, <creaky>). These medium-level attributes are in turn described by a set of low-level acoustic attributes defining the perceptual correlates of the sound (e.g. <spectral tilt>, <shimmer>, <jitter>). The low-level acoustic attributes correspond to the acoustic controls that the extended MBROLA synthesizer can render through the sound processing procedure described above. This descriptive scheme has been implemented within FESTIVAL as a set of mappings between high-level and low-level descriptors. The implementation includes the use of envelope generators to produce time curves of each parameter.

In order to check and evaluate, by direct low-level manual/graphic instructions, various multi level emotional facial configurations we developed an EmotionPlayer, which was strongly inspired by the EmotionDisc of Zsafia Ruttkay [Ruttkay et al., 2003]. It is designed for a useful immediate feedback, as exemplified in Fig. 2.

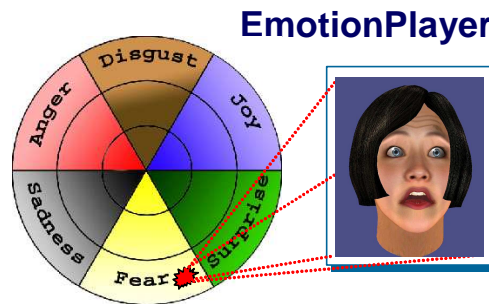


Figure 2. Emotion Player: clicking on three-level intensity (low, mid, high) emotional disc, an emotional configuration (i.e. high -fear) is activated.

## 5. CONCLUSION

LUCIA is an MPEG-4 standard FAPs driven OpenGL framework which provides several common facilities needed to create a real-time facial animation application. It has high quality 3D model and a fine coarticulation model, which is automatically trained by real data, used to animate the face.

The modified coarticulatory model is able to reproduce quite precisely the true cinematic movements of the articulatory parameters. The mean error between real and simulated trajectories for the whole set of parameters is, in fact, lower than 0.3 mm.

Labial movements implemented with the new modified model are quite natural and convincing especially in the production of bilabials and labiodentals and remain coherent and robust to speech rate variations.

The overall quality and user acceptability of LUCIA talking head has to be perceptually evaluated [Massaro, 1997 - Costantini et al., 2004] by a complete set of test experiments, and the new model has to be trained and validated in asymmetric contexts. Moreover, emotions and the behavior of other articulators, such as tongue for example, have to be analyzed and modeled for a better realistic implementation.

A new WebGL implementation of the avatar is currently in progress to exploit new possibilities that arise from the integration of LUCIA in the internet websites.

## REFERENCES

- Massaro, D.W. et al., 2000. Developing and Evaluating Conversational Agents. *Embodied Conversational Agents*, MIT Press, Cambridge, USA, pp. 287-318.
- Bregler, C. et al., 1997. Video Rewrite: Driving Visual Speech with Audio. *Proceedings of SIGGRAPH '97*, pp. 353-360.
- Terzopoulos, D. et al., 1995. Realistic Face Modeling for Animation. *Proceedings of SIGGRAPH '95*, pp. 55-62.
- Vatikiotis-Bateson, E. et al., 1996. Physiology-Based Synthesis of Audiovisual Speech. *Proceedings of 4th Speech Production Seminar: Models and Data*, pp. 241-244.

- Beskow, J., 1995. Rule-Based Visual Speech Synthesis. *Proceedings of Eurospeech '95*, Madrid, Spain, pp.299-302.
- LeGoff, B. and Benoit, C., 1996. A text-to-audiovisual speech synthesizer for French. *Proceedings of ICSLP '96*, Philadelphia, USA, pp. 2163-2166.
- Farnetani, E. and Recasens, D., 1999. Co-articulation Models in Recent Speech Production Theories. *Co-articulation in Speech Production*. Cambridge University Press, Cambridge, USA
- Bladon, R.A. and Al-Bamerni, A., 1976. Co-articulation resistance in English. *Journal of Phonetics*. No.4, pp. 135-150
- Cosi, P. and Perin, G., 2002. Labial Co-articulation Modeling for Realistic Facial Animation. *Proceedings of ICMI '02*, Pittsburgh, USA, pp. 505-510.
- Fusaro, A. et al., 2003. LUCIA a New Italian Talking-Head Based on a Modified Cohen-Massaro's Labial Co-articulation Model. *Proceedings of Eurospeech 2003*, Geneva, Switzerland, vol. III, pp. 2269-2272.
- Ferrigno, G. and Pedotti, A., 1985. ELITE: A Digital Dedicated Hardware System for Movement Analysis via Real-Time TV Signal Processing. *In IEEE Transactions on Biomedical Engineering*, BME-32, pp. 943-950.
- Tesser, F. et al., 2001. Festival Speaks Italian!. *Proceedings of Eurospeech 2001*, Aalborg, Denmark, pp. 509-512.
- Cosi, P., and Magno-Caldognetto, E., 1996. Lip and Jaw Movements for Vowels and Consonants: Spatio-Temporal Characteristics and Bimodal Recognition Applications. *In Speechreading by Humans and Machine: Models, Systems and Applications*. NATO ASI Series, Series F: Computer and Systems Sciences, vol. 150, pp. 291-313.
- Zmarich, C. et al., 1997. Italian Consonantal Visemes: Relationships Between Spatial/temporal Articulatory Characteristics and Coproduced Acoustic Signal. *Proceedings of AVSP '97*, Rhodes, Greece, pp. 5-8.
- Magno-Caldognetto, E. et al, 1998. Statistical Definition of Visual Information for Italian Vowels and Consonants. *Proceedings of AVSP '98*, Terrigal, Austria, pp. 135-140.
- Boersma, P., 1996. PRAAT, a system for doing phonetics by computer. *In Glot International*, No. 5, pp. 341-345.
- Magno Caldognetto, E. et al., 2003. Coproduction of Speech and Emotions: Visual and Acoustic Modifications of Some Phonetic Labial Tar-gets. *Proceedings of AVSP 2003*, St Jorjioz, France, pp. 209-214.
- Drioli, C. et al., 2003. Emotions and Voice Quality: Experiments with Sinusoidal Modeling. *Proceedings of Voqual 2003*, Geneva, Switzerland, pp. 127-132.
- Tisato, G. et al, 2005. INTERFACE: a New Tool for Building Emotive/Expressive Talking Heads. *Proceedings of INTERSPEECH 2005*, Lisbon Portugal, pp. 781-784.
- Hartman, J. and Wernecke, J., 1996. *The VRML Handbook*. Addison Wessley.
- Tesser, F. et al., 2004. Prosodic Data-Driven Modelling of Narrative Style in FESTIVAL TTS. *Proceedings of 5th ISCA Speech Synthesis Workshop*, Pittsburgh, USA.
- Tesser, F. et al, 2005. Emotional Festival-Mbrola TTS Synthesis. *Proceedings of INTERSPEECH 2005*, Lisbon, Portugal, pp. 505-508.
- Drioli, C. et al., 2005. Control of Voice Quality for Emotional Speech Synthesis. *Proceedings of AISV 2004*, Padova, Italy, pp. 789-798.
- Nicolao, M. et al., 2006. GMM modelling of voice quality for FESTIVAL-MBROLA emotive TTS synthesis. *Proceedings of INTERSPEECH 2006*, Pittsburgh, USA, pp. 1794-1797.
- Cosi, P. et al., 2004. Data-Driven Tools for Designing Talking Heads Exploiting Emotional Attitudes. *Proceedings of Tutorial and Research Workshop "Affective Dialogue Systems"*, Kloster Irsee, Germany, pp. 101-112.
- Magno-Caldognetto, E. et al., 2004. Visual and acoustic modifications of phonetic labial targets in emotive speech: Effects of the co-production of speech and emotions. *In Journal of Speech Communication*, vol. 44, pp. 173-185.
- Cavicchio, F. et al, 2004. Modification of the Speech Articulatory Characteristics in the Emotive Speech. *Proceedings of Tutorial and Research Workshop "Affective Dialogue Systems"*, Kloster Irsee, Germany, pp. 233--239.
- Pelachaud, C. et al., 2004. APML, a Mark-up Language for Believable Behavior Generation. *In Life-Like Characters*, Springer, pp. 65-85.
- Ruttkay, Z. et al., 2003. Emotion Disc and Emotion Squares: tools to explore the facial expression space. *In Computer Graphics Forum*, No. 22, pp. 49-53
- Massaro D.W., 1997: Perceiving Talking Faces. *Speech Perception to a Behavioral Principle*. MIT Press, Cambridge, USA
- Costantini, E. et al., 2004. Evaluation of Synthetic Faces: Human Recognition of Emotional Facial Displays. *Proceedings of Tutorial and Research Workshop "Affective Dialogue Systems"*, Kloster Irsee, Germany, pp. 276-287.

# ARABIC USABILITY GUIDELINES

Perihan Elbaz

*Computer Sciences and Information Department,  
Institute of Statistical and Research Studies (Egypt)*

## ABSTRACT

Usable interfaces are a result of focus on understanding user's culture and subcultures. The aim of this paper is to set the scene for research into extract Arabic usability guidelines. This paper proposes extracting of the Arabic usability guidelines for Arab culture and subcultures based on the issues gleaned from user acceptance test (UAT) reports. After applying the proposed guidelines on Web application projects that serve different sectors. The application of the guidelines led to a decrease of UAT bugs by 95.25%, and an increase of customer confidence and satisfaction. Systems' usability also increased as evidenced by the results obtained from usability tests, and user interface analyses (Elbaz, 2007, Elbaz, 2010).

## KEYWORDS

Culture, Usability, User-Centered Design, and Arabic Web Pages Guidelines.

## 1. INTRODUCTION

It is better to make the system aesthetic appeal for the end user. Making products and systems easier to use, will reduce the cognitive load on users. The most obvious way of enhancing performance is to design systems that will reduce the amount of cognitive resources (Mayhew, 1992, Baecker and Buxton, 1987). Thus, in recent years, usability has become more than a concept. In many organizations, usability is an important part of software development. Moreover, more and more companies are discovering that usability is not only good for users, but also good for business, for developers and for organization members (Weinschenk, 2005).

The importance of guidelines was first revealed during the eighties (Reed et al., 1999) when the use of the computer dramatically increased in the work place: more computer-based systems were used by more users, who were not necessarily expert, for more interactive tasks, possibly new or unusual. The advent of Web sites and Web-based applications raised the amount of sources containing guidelines for the Web User Interfaces (UIs). Among others were published seminal books such as (Nielsen, 2000), style guides produced by individuals such as (Lynch and Horton, 1999) or by organizations such as (Brown, 1999), standards such as (ISO, 1999), sets of design rules, and lists of principles. A wide variety of usability guidelines exist, and have been established by different authors (Rivlin et al., 1990, Bevan et al., 1991, Preece et al., 1994, Rigden, 1999, Nielsen and Norman, 2000, Ivory, 2000, Lynch and Horton, 2009, Ivory and Hearst, 2002, Lengel, 2002, Nielsen, 2002, Rohn et al., 2002, Graham, 2002). Each one of them focuses on how to satisfy users by presenting a usable Web design. These guidelines address a wide range of Web page design issues starting from the browser's title to the detail of Web page features (Ivory, 2000).

Using usability sources is not straightforward for developers and evaluators, primarily for some of the following reasons (Reed et al., 1999): Usability remains a quality factor of user interfaces. It is hard to select guidelines appropriate to a particular context of use. Thus, the paper aim is to give attention of Arabic usability guidelines as it is the first focus into details of Arabic language within Web sites and applications.

## 2. METHODOLOGY

In order to extract the Arabic usability guidelines that we aim for, we have carried out the following steps: Surveying the literature for related research work. Bottom-up approach is used for a case study (Elbaz, 2010).

### 3. A CASE STUDY

#### 3.1 Data Collection

We have collected UAT bugs data records from the thirteen projects developed by four companies. The projects included private, governmental and E- Government sectors. Total UAT issues 2164 and total number of usability issues 715 with percentage 33%. Table 1 below illustrates the usability guidelines elements actually used and the relative frequencies of their occurrence in the UAT reports, expressed as percentages (Elbaz, 2010).

Table 1. Usability Guidelines Elements Percentage

Usability Guidelines Elements	Percentage	Usability Guidelines Elements	Percentage
Content Organization	32%	Graphics	2.1%
Navigation	25.2%	Text appearance	2%
Labels	15.9%	Scrolling and paging	1.1%
Page Layouts	12.6%	Color	1.4%
Search	5.9%	Alignments	1.8%

Then determine usability issues that are influenced by objective culture variables according to Galdo (Del Galdo, 1996). Formats (Time-Date-Telephone) 1.3%, Directions 2.2%, Translations 1.1% and Currency .3%.

### 4. EXTRACTING ARABIC USABILITY GUIDELINES

The idea is to write almost complete guidelines suitable for Arab countries Web sites/ applications based on the extracted usability issues from UAT reports (Elbaz, 2010) and any other RTL scripts.

#### 4.1 The Home Page Structure

Announce major changes into the home page. These announcements should be decorative and recognized from first look in the home page. Most Arab countries specially, Kuwait mapping to their kind of their life style wanting every things in decorative way. The latter statement comes from a case study. For RTL (right to left direction) languages such as Arabic, most important information should be placed in the right. Right panels should be clear to the user to get their purpose. Then place contents in the middle and left. (Elbaz, 2010)

#### 4.2 Page Layout

For Arab countries, put banner that includes logged user, department, date/time and other options can be used to set focus on them. Choose appropriate line height. Most of Arabic characters have tails under the straight line. Thus, small line length makes words with narrow spaces. To satisfy Arab countries for making headers for example very large in height as they can take 1/3 page length includes huge graphics / decorations / flashes. Let it in the home page and in the navigation pages sit option to auto hide header as same as windows task bar. (Elbaz, 2010)

#### 4.3 Navigation

Avoid popup window or open new browser windows that confuse users. Can be resolved using popup Ajax control or make a layer over the main page and let the main page not accessible to set focus on the new popup or layer. Use tabs in proper way (Bailey et al., 2006, Kim, 1998). As most of Arab business users dislike Tabs. If navigation needs to use Tabs, use tabs that are more likely to find and click appropriately on tabs that look like real-world tabs (Elbaz, 2010) based on heuristic experience.

#### **4.4 Scrolling and Paging**

Use paging more than scrolling. Paging can be for 'Tables' or in the 'Data Grid control' in dot net technology as paging is built in. Divide paging 10 by 10 to display more data and when using 'Data Grid' within long page let paging size 5 (Elbaz, 2010).

#### **4.5 Links**

Make sure that link names are consistence with their target. Arab countries can use more than two words to clearly what actually the target from link (Elbaz, 2010).

#### **4.6 Text Appearance**

Do not use ornamental fonts as they are hard to read. To satisfy Arab countries, use it in the banners inside flash for example. Apply large font in displaying names (Elbaz, 2010).

#### **4.7 Screen-Based Controls**

Minimize the amount of information entered by users. In most cases, Arab countries want all information in the same page, in such case use "tabs, show/hide" of some parts to let user focus into the entry fields. Use (\*), rather than bolding the labels with display message in the top of the page indicates that all fields with (\*) remark are mandatory before going through data entry then discover that user should fill those items. Provide auto-tabbing functionality. Provide maximum length for each control for data entry fields. For names, provide 50+ characters (according to Arabic names specially, in KSA) (Elbaz, 2010).

#### **4.8 Graphics, Images and Multimedia**

Do not use icons represent people, animals, or religious symbols (Elbaz, 2010, Elbaz, 2007, Elbaz, 2009). Use impressive graphics and flashes (Elbaz, 2010, Elbaz, 2007, Elbaz, 2009). Display errors in proper format. In Kuwait project for example, use error messages in separate page with colorful themes and decorative things (Elbaz, 2010).

#### **4.9 Content Organization**

Group all related information and functions using heading or "field set" in order to decrease time spent searching or scanning. Separate between males and females in forums (Elbaz, 2010, Elbaz, 2007, Elbaz, 2009).

#### **4.10 Search**

Add more than one parameter to reduce search scope to speed up search result. Support 'auto complete' to speed up data entry. Trim spaces from left and right (Elbaz, 2010).

Thus, by the ending of this section, most of the Arabic usability guidelines extracted from UAT reports is presented.

### **5. EXTRACTED USABILITY GUIDELINES VALIDATION**

The proposed guidelines were validated through empirical trials through two projects. We will focus on one project. This project was located in KSA (Kingdom of Saudi Arabia). The project was related to a banking sector. The number of project's business processes is around 80 processes. A software development company stared the implementation of the first three business processes out of eighty. Feedback on the developed was not as expected from the customer's and the end users. The software company that applied the proposed

framework took over the developed processes in terms of eliminating bugs. The management with the customer decided to rebuild the three processes from scratch due to the huge number of bugs and develop another seven processes (Elbaz, 2010, Elbaz, 2007). The User centered Design (UCD) department within local software company located in Egypt started implementing the screens using the extracted Arabic usability guidelines. From the UAT reports total UAT issues before were 2333 and after applying issues were 111 only. The UAT bugs are decreased with this percentage (95.25%). There was a decrease in the development time by (50%). In one of the meeting minutes, a customer said the following (after translate it into English): “After using the application, the IT and business users were satisfied with the application and how it is very easy to use. They thanked all the development and the UCD team for their specialization and the professional work that represents worldwide standards in the application”. This feedback is based on the usability tests and the user interface analysis by an HCI expert. The user interface analysis is based on human factors principles as shown in table 2 (Elbaz, 2010, Elbaz, 2007).

Table 2. User Interface Analysis

Metrics Class	Metrics	Evaluation (Before)	Evaluation (After)
Learnability	Predictability	3- Acceptable	3-Acceptable
	Synthesizability	4- Fair	4- Fair
	Familiarity	4- Fair	4- Fair
	Generalizability	3- Acceptable	4- Fair
	Consistency	3- Acceptable	4- Fair
Flexibility	Dialogue Initiative	4- Fair	4- Fair
	Multi-Threading	2- Poor	3-Acceptable
	Task Migratability	2- Poor	3-Acceptable
	Substitutivity	2- Poor	3- Acceptable
Flexibility	Customizability	2- Poor	3-Acceptable
Robustness	Observability	4- Fair	4- Fair
	Recoverability	3- Acceptable	4- Fair
	Responsiveness	3- Acceptable	4- Fair
	Task Conformance	4- Fair	4- Fair
Other	Use knowledge of real world	3- Acceptable	4- Fair
	Task Simplification	3- Acceptable	4- Fair

## 6. CONCLUSIONS

Using of the Arabic usability guidelines will increase the usability quality of Arab countries' projects. After applying the proposed guidelines, there was a decrease in the UAT bugs by 95% on average. Customer satisfaction and confidence are increased. There was an Increase in the system usability. The research reported herein clearly points into the influence of culture on extracting Arabic usability guidelines. The ways to combat such effects is important and necessary to allow designers to design a positive relationship between the project and the human, rather than a negative one. In addition to, the utility of the UCD approach in software localization efforts, which suggests that further research into extracting guidelines for the Arabic-speaking culture (Elbaz, 2010, Elbaz, 2007, Elbaz, 2009).

## ACKNOWLEDGMENT

All praise be to Allah for supporting me in each step in this research. The completion of this paper would not have been possible without the help and support of many individuals to whom I owe a great deal of thanks.

## REFERENCES

- BAECKER, R. M. & BUXTON, W. A. S. (1987) *Readings in human-computer interaction: A multidisciplinary approach*, Morgan Kaufmann Publishers, Los Altos, CA.
- BAILEY, R. W., BARNUM, C., BOSLEY, J., CHAPARRO, B., DUMAS, J., IVORY, M. Y., JOHN, B., MILLER-JACOBS, H. & KOYANI, S. J. (2006) *Research-Based Web Design & Usability Guidelines*.
- BEVAN, N., KIRAKOWSKI, J. & MAISSEL, J. (1991) what is usability. *Proceedings of the 4th International Conference on HCI*. Stuttgart.
- BROWN, P. (1999) *Java Look and Feel Design Guidelines*, Addison-Wesley.
- DEL GALDO, E. M. (1996) Culture and design. *International users interface*. John Wiley & Sons, Inc.
- ELBAZ, P., GALAL-EDEEN, GALAL HASSAN (2007) The Influence Of Culture On System Usability. IN M. H. EL-WANY, A. B. E. (Ed.) *Proceedings of the 37th International Conference on Computers and Industrial Engineering*. Alexandria, Egypt.
- ELBAZ, P., GALAL-EDEEN, GALAL HASSAN (2009) Adaptation of User-Centered Design Methods. IN S., T. M. F. A. A. E.-W. (Ed.) *Fourth International Conference on Intelligent Computing & Information Systems*. Cairo-Egypt.
- ELBAZ, P., GALAL-EDEEN, GALAL HASSAN (2010) The Effect of Culture on User-Centered Design Methods and Phases. IN VENESSA EVERS, J. A.-N., ELISA DEL GALDO (Ed.) *The 9th International Workshop on Internationalisation of Products and Systems (IWIPS)*. London United Kingdom, and hosted by Thames Valley University.
- GRAHAM, I. (2002) *A pattern language for web usability*, Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA.
- ISO (1999) Human-Centered Design Processes for Interactive Systems. *International Standard; ISO Standard 13407*. Geneva, Switzerland.
- IVORY, M. Y. (2000) Web TANGO: towards automated comparison of information-centric web site designs. *Human Factors in Computing Systems*. The Hague, The Netherlands, ACM New York, NY, USA.
- IVORY, M. Y. & HEARST, M. A. (2002) Statistical profiles of highly-rated web sites. *Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves*. Minneapolis, Minnesota, USA ACM New York, NY, USA.
- KIM, J. (1998) An empirical study of navigation aids in customer interfaces. *Behavior and Information Technology*, 18, pp. 213-224.
- LENGEL, J. G. (2002) *The Web wizard's guide to Web Design*, Pearson Education, Inc.
- LYNCH, P. J. & HORTON, S. (1999) *Web style guide*, Yale University Press New Haven.
- LYNCH, P. J. & HORTON, S. (2009) *Web style guide: Basic design principles for creating Web sites*, Yale University Press.
- MAYHEW, D. J. (1992) *Principles and guidelines in software user interface design*, Prentice-Hall, Inc. Upper Saddle River, NJ, USA.
- NIELSEN, J. (2000) *Designing web usability*, New Riders
- NIELSEN, J. & NORMAN, D. A. (2000) Web-site usability: usability on the web isn't a luxury. *Information Week Online*, 14, pp. 65-73.
- NIELSEN, J. T. (2002) *Homepage Usability: 50 Websites Deconstructed*, New Riders Publishing, Indianapolis.
- PREECE, J., ROGERS, Y. & BENYON, D. (1994) *Human-computer interaction*, Addison-Wesley.
- REED, P., HOLDAWAY, K., ISENSEE, S., BUIE, E., FOX, J., WILLIAMS, J. & LUND, A. (1999) User interface guidelines and standards: progress, issues, and prospects. *Interacting with Computers*, 12, pp. 119-142.
- RIGDEN, C. (1999) The Eye of the Beholder'-Designing for Colour-Blind Users. *Telecommunications Engineering*, 17, pp. 291-295.
- RIVLIN, C., LEWIS, R., DAVIES-COOPER, R. & ECONOMIC AND SOCIAL RESEARCH, C. (1990) *Guidelines for screen design*, Blackwell Scientific Publications.
- ROHN, J. A., SPOOL, J., EKTARE, M., KOYANI, S., MULLER, M. & REDISH, J. G. (2002) Usability in practice: alternatives to formative evaluations-evolution and revolution. ACM New York, NY, USA.
- WEINSCHENK, S. (2005) Usability a business case. *White Paper Human Factors International*.

# COMMUNICATION SUPPORT BETWEEN PATIENTS DURING TREATMENT VIA USE OF MEMORIES

Takahiro Kihara<sup>1</sup> and Yoshio Nakatani<sup>2</sup>

<sup>1</sup>*Graduate School of Science and Engineering*

<sup>2</sup>*College of Information Science and Engineering  
Ritsumeikan University*

*1-1-1 Noji-higashi, Kusatsu, Shiga, Japan*

## ABSTRACT

Patients in long-term treatment due to accidents, disasters, or disease often feel lonely and have more of a sense of loss in inorganic spaces such as sickrooms than people being currently treated. This study therefore involves a proposition of a framework that supports the chance for those long-term patients to talk about each other's fond memories when they meet during their treatment. To invoke communication with others recuperating by passing on information to each other via the nurses in charge, the system offers fond memories of a common topic and thus offers information that enables the chance of making a self-introduction for the user. The expectation is that mutual understanding can be deepened via that support between patients, patients and nurses, patients and other patients, and patients and doctors. Moreover, it could support the creation of better interpersonal relationships within hospitals.

## KEYWORDS

Fond Memory Communication, Patient's chats, Private self, Public self

## 1. INTRODUCTION

People have various worries in their life. They often suffer from loneliness and a sense of loss. Moreover, their hopes in life can be lost when placed in a situation in which their living conditions have drastically changed, with some people even feeling, "They wish to die" or "They were deserted by their families". Communication can provide relief from such negative feelings. Several studies in this field have focused on communication between doctors and patients to date. However, very few attempts have been made to study communications between patients and between nurses and patients.

A framework that enables communication to take place between patients is an indispensable element during long periods of recuperation. Patients have to spend a lot of time in the sickrooms made up of inorganic, unknown spaces when they are hospitalized. It is thought that the patients conversing would be effective in relieving their loneliness.

Patients often express dissatisfaction that their stories cannot be easily heard because the doctors and nurses are very busy. Conversely however, doctors and nurses often wish they could listen to as many patients' stories as possible, and hence actually wants to hear them. As proposed in this study it is thought that inducing conversation between patients would have the advantage of enabling doctors and nurses to easily participate in the conversation, too.

### 1.1 About the "Narrative" Approach

There is a concept in medicine that is based on conversations taking place between doctors and patients as an approach at medical treatment sites. The phrase "the patient's story", as used here, chiefly indicates "the patient's experience with struggling with a disease." "DIPEX Japan" and "Life Palette" include "experiences with health and sickness" in a data base as part of a relational approach. They are expected to be of use in social support anyone distressed with the same sickness because a variety of experience from struggling with



a disease can be "set wisdom." Moreover, the effect of the catharsis as depuration of the mind can be expected from talking to other people [1].

## 1.2 Communication via Personal Memories

"Personal memories" can be used to not only recall things yourself but also to communicate with others. Yamashita and Nojima named the communication function of personal memories "Memories communication [2]." Approaches exist that use photographs to trigger recollection of personal memories, for example via "Photologue Viewer" and "Electronic mini album [3]."

## 2. METHOD

Patient's talking to each other in this study doesn't mean the narrow sense of about "the recorded experience of struggling with a disease" and instead also past "personal memories."

There has been no attempt to date to directly connect patients, although there are a lot of approaches that center on a patient.

In local spaces such as hospitals and facilities there are typically a number of people suffering from the same sickness, facing the same problems, and basically in the same situation. In this study a system is proposed to enable those people when they meet in hallways or sickrooms to present some common personal memories in thus inducing the opportunity to converse with each other (Figure 1).

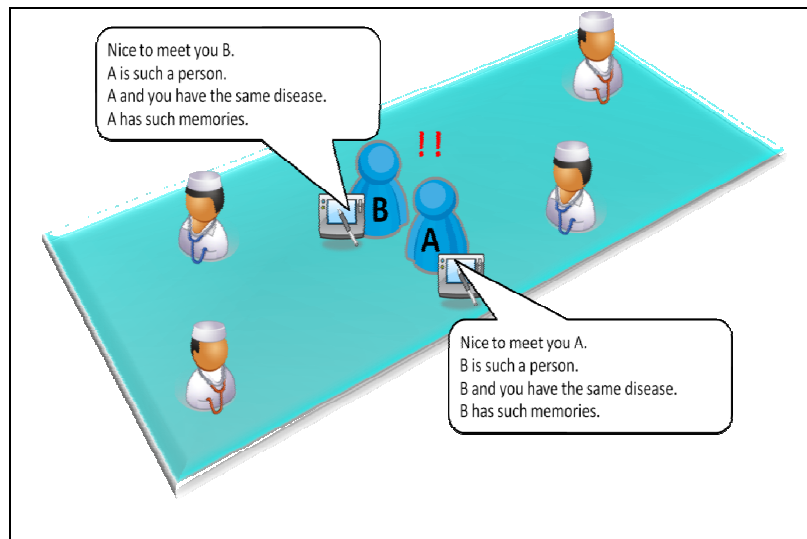


Figure 1. Imagined system utilization scenario

### 2.1 Analysis of First Conversation

This system presents individual information as a self-introduction and at the same time a personal memory. Japanese typically initiate first meetings with a formal exchange of a formulaic greeting. Especially when they first meet someone publically they only talk about their work, hobbies, and home life. Their private lives, which include past experiences and feelings, are not talked about with others until they trust the other party [4]. Because of this the system casually presents the following four points when they introduce themselves.

1. Greeting of "Nice to meet you".
2. How many times have you met the other party?
3. When did you meet?
4. What do you have in common with the other party?

The expectation is that this will then result in users gradually acquiring the chance to converse in a trustworthy manner with the other person.

## 2.2 Mode of Expression via Accompanying Metadata

It is said that humans acquire 80% to 90% of information via the sense of sight, thus leading to the consideration that looking back on personal memories via viewing photographs and mementos is an effective approach to take [5]. In this study personal memories are invoked by focusing on “date and time the picture was taken”, “place the picture was taken” and “content of events” three minimal pieces of information using a variety of metadata associated with a picture.

## 3. PROTOTYPE SYSTEM

A system called a "Memory exchange system" that was proposed in Chapter 2 was constructed using a Windows PC, and the prototype system explained in this chapter.

For example conversations between patients passing each other are invoked in the following processes.

1. "Memory exchange system" in which wireless applications are distributed and attached to patients in hospitals beforehand. All the addresses of each wireless application are recorded on the equipment.
2. Patients equipped with wireless applications recognize each other via wireless telecommunication.
3. If the detected wireless address has been registered in the database process 4 is executed. It returns to 2 if not registered.
4. Metadata on your own fond memories and the other party's are compared, and commonality then set according to the linked tags. Common memories of the highest level are then prioritized.

### 3.1 Content of System

The Bluetooth equipment has individual addresses. The system enables information on other parties to be acquired. The Bluetooth address of an individual device is automatically detected if registered in the database as an ID. The name of the disease from which the user suffers is also registered. Moreover, some people do not want others to know the disease they have. Figure 2 shows the configuration of the system.

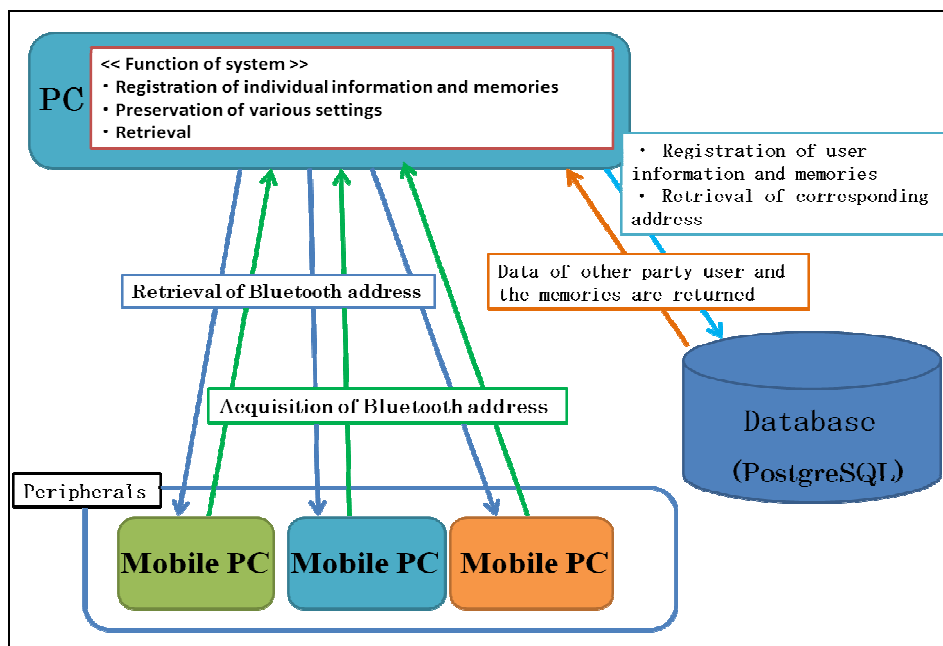


Figure 2. Figure of system configuration

Moreover, the results of execution are shown in Figure 3 when the corresponding address is in the database.

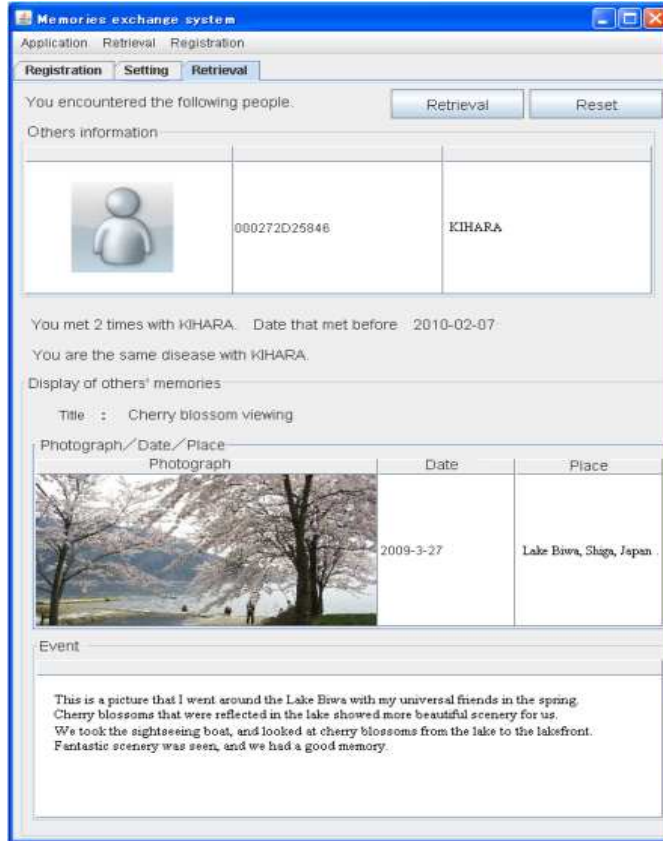


Figure 3. Example of system screen

## 4. EVALUATION

Once the prototype system was complete the opinions of five Japanese occupational therapists at the "Medical corporation Kofu association of the Munakata hospital rehabilitation department in Fukuoka Prefecture Munakata City" were requested. The special aspects of "Utility of the system on-site" and "Problems when the system was actually used onsite" were then evaluated.

### 4.1 Evaluation by Specialist

The background of the research, purpose of the system utilization, and operating method were explained in a demonstration of the system before the evaluation. Opinions, needed improvements, and future work concerning this system were then requested.

The session took place at Munakata City, Fukuoka, Munakata hospital, JAPAN on January 28, 2010.

#### 4.1.1 About Activation of a Medical Treatment Site

If an environment in which the system is available activation of a medical treatment site can be expected to result from the system. Patients can often be uneasy, timid, and have to depend on others when they are sick. If they can find something in common with someone else it can be very reassuring to them. Moreover, because knowing another party's fond memories can be a chance to get to know the other party better a good relationship can be built.

### 4.1.2 About Patient's Communications

The system could be very useful to young people accustomed to using personal computers. However, it could prove difficult for anyone not at all accustomed to them before using the system, and therefore cannot be said to be useful for all patients being treated. However, it can be expected the very useful as a communication tool for doctors, nurses, and patients during treatment if made available for use in hospitals.

### 4.1.3 Improvement of System

Introducing the system at the start of hospitalization can be assumed to be a heavy burden on users and therefore it was pointed out that it should not be introduced when patients are first hospitalized and instead after they have become accustomed to hospitalization and making acquaintances. Moreover, there is a possibility that it would originate in further recovery if introduced when during convalescence.

## 5. CONCLUSION

Users may strongly resist using such a system within medical treatment sites where personal computers are not very frequently used. The convenience of the system therefore needs to be re-examined.

The effectiveness of the system could be confirmed in that one's own memories could be arranged and the chance to converse enhanced through use of it. However, the effectiveness of introducing the system when first hospitalized needs to be re-examined, as indicated in the comment made by the specialist.

Based on above future work will include the following. The opinion was aired that "It would be convenient if a function where only health care professionals were able to display part of the clinical records if they also had this system were provided" by an occupational therapist. The clinical history of a patient during treatment and their physical condition are examples. This could also include past treatment and today's schedule in the case of rehabilitation. It can also be expected that the system would lead to reduced incidents if it was capable of displaying information on the diseases of patients during treatment.

## ACKNOWLEDGEMENTS

Special thanks go to Mami Takada, Sayaka Imaizumi, Ryojiro Fukada, Keiji Fujita, and Keiko Ogawa whose opinions and information also helped a lot throughout the production of this study.

## REFERENCES

- [1] Matsubara, T., Niregi, M., Sawada, T., Miyagi, M., 2005, *Encyclopedia for psychological care*. BAIFUKAN Publishers, Tokyo, JAPAN (in Japanese)
- [2] Yamashita, K., Nojima, H., 2001, *Portable Personal Memory for Communication*. Human Interface Symposium 2001. Sendai, Japan, pp. 503–506. (in Japanese)
- [3] Yamashita, K., Nojima, H., 2002, *An Arrangement Method of Digital Photographs Based on Importance of Photographs*. *The 19th Annual Meeting of the Japanese Cognitive Science Society*, Ishikawa, Japan, pp. 256–257. (in Japanese)
- [4] CHANG, Y. 2006, A Contrastive Study of Initial Interactions between Taiwanese and Japanese Female University Students. *Journal of the Humanities in Japan*, Vol. 9, pp 223-233.
- [5] Nakatani, Y., 2005, Support System for Reconstructing Fond Memories Lost in Disasters. *11th International Conference on Human-Computer Interaction 2005*.

# SMART BEER COASTER

Leonardo Rocha, Augusto Lohmann, André Braz, Breno Bitarello and Fernando Reiszal

*ESDI-UERJ*

*Evaristo da Veiga 95, Rio de Janeiro, RJ, Brazil*

## ABSTRACT

The main objective of this paper is to describe the different stages of research and development of an electronic device based on the use of open source technology which is the result of the observation of the consumption and sales habits of 600ml beer bottles in Brazil, as well as the relationship between bars/restaurants and their customers. This project's goal is to optimize this business, being both commercially attractive and capable of improving its consumption experience. Therefore, the prototype here presented is expected: 1) to warn the staff that the last bottle served to a customer is getting empty, allowing its immediate replacement and consequently reaching a higher level of client satisfaction; 2) to increase the efficiency and profit margin of the bar/restaurant by selling more beer; 3) to add value to the bar/restaurant and the brands associated to the device as a product, due to its technological and innovative nature.

## KEYWORDS

Design, consumption, beer, interface, communication

## 1. INTRODUCTION

The increase in the development of open source technologies and the price reduction of electronic components have facilitated the prototyping process of gadgets in the most varied areas of knowledge.

Aware of the actual relationship between bars/restaurants and their customers, we noticed the possibility of using technological resources to optimize it. Based on a study on the service of selling and consuming 600ml beer bottles, we can point out some conditions and characteristics of this activity flow that could be modified and optimized with the aid of smart and ubiquitous technologies, according to the list below:

- To request a new beer after consuming the previous one, the customer needs to realize that their beer has finished and calls the waiter with a gesture or word. In many cases, the waiter is not immediately within the customer's scope of vision or not aware he has been called, which results in a delay to get more beer. In a greater scale, with a large number of customers the hypotheses is that such waiting to drink tends to represent a potential loss of consumption by customers and sales by the bar/restaurant;
- The bar/restaurant needs, especially at peak times, a greater number of waiters resulting in their circulation around the work place in an irregular way. On the other hand, an organized and well planned circulation allows a coverage of the whole useful area with less staff and less waste of time;
- The staff, since they do not have precise information on which tables need service, increase their stress load and physical effort in the process of constantly supervising customers' consumption, which, in the long run, may result in a reduction of satisfaction and service capacity.

## 2. PURPOSES

Our first purpose is the research on the technical matters for the development of the intended prototype, in addition to the use flow and its requirements, so that it may increase efficiency in the beer sales service.

With the prototype correctly implemented and configured for the service of selling 600ml beer bottles from Ambev, a Anheuser-Busch InBev company, the focus is to provide a system that warns the staff as the container becomes closer to the "empty" status and requires its replacement, giving more satisfaction to customers and making the work flow more efficient, which tends to sell more beer in a given amount of time.

Due to the fact that it has been entirely developed by using open source technologies, the prototype presents a low-cost solution, with an initial investment that may be easily recovered as revenue increases and, consequently, as the profit margin of the bar/restaurant increases too, considering the optimization of the service flow provided by the use of the device.

Likewise, one of the goals to be achieved with the use of the prototype is the optimization of the staff shift schedule, allowing the bar/restaurant to provide a better service with fewer employees.

Finally, there is a concern to develop a device in which its attributes – such as its design as well as its innovative and technological characteristics – become themselves an attraction to the customers, a user friendly item that could be perceived as a product that adds value to the bar/restaurant's service in general.

### 3. TECHNICAL SPECIFICATIONS

The solution adopted in the project consists in two essential elements: coaster and receiver. The coaster, which provides support to the container (beer bottle) and records the pressure it exerts on its surface (contact area), forwards data via radio to the receiver from time to time – which will be connected to a computer with an application and graph interface to the intended goal – in such a way that it is possible to follow up the status of the container and its prospective replacement. The information obtained in this process may integrate a logistic and automation system to control the consumption of the table where the coaster is located, as well as the stock of products available in the bar/restaurant.

In order to gather and send to the receiver the data containing the pressure exerted by the container, the coaster must have the following items:

1. Square Force Sensitive Resistor – code SEN-09376, price: U\$7,90 at SparkFun Electronics;
2. Arduino Fio – code DEV-09712, price: U\$24,95 at SparkFun Electronics;
3. Polymer Lithium Ion Battery - 1000mAh – code PRT-00339, price: U\$11,95 at SparkFun Electronics;
4. XBee 1mW Chip Antenna – code WRL-08664, price: U\$22,95 at SparkFun Electronics.

For the receiver, the following items were used:

1. XBee Explorer Dongle – code WRL-09819, price: U\$24,95 at SparkFun Electronics;
2. XBee Pro 60mW Chip Antenna – code WRL-08690, price: U\$37,95 at SparkFun Electronics.

For economic reasons, it has been used the minimum number of parts in the prototyping, both for the coaster and the receiver, in such a way to make the solution feasible within the current market scenario. Besides, there is a clear concern in relation to the physical space occupied by the devices comprising the coaster because its dimensions cannot exceed those of the typical cork coaster. Therefore, we decided to use the Arduino Fio, which is compact, being equipped with a socket for XBee and connector for Li-Po battery. However, its mini-USB connector serves only for the purposes of recharging the battery.

The programming of the coaster is made remotely from the XBee Explorer Dongle that integrates the receiver, being unnecessary to connect the Arduino Fio to a computer. However, to recharge the battery that feeds the board, it is necessary to connect it to a USB power source. Replacing the discharged battery by a charged one would also be a possible solution.

Although the container, that shapes the coaster and accommodates the four items it contains, has not been built in this project, we estimate that its dimensions are approximately 90 x 90 x 22mm (length x width x height), being different only in height in relation to the common coasters.

As to the receiver, the option fell upon XBee Explorer Dongle, which has a connector/USB feeder, and a socket for XBee; the latter, in turn, was occupied by a 60mW XBee Pro to ensure communication with the coaster 1mW XBee, since it is not possible to know the oscillation of signal in view of innumerable physical barriers (walls, pillars, people, etc) between both XBees, at distances higher than 15m.

The size and format resulting from the combination described above cause the receiver to have an appearance similar to that of a pen drive, although, just like the coaster, an enclosure has not been built to protect the equipment and make its handling easy.

#### 4. PROTOTYPE: TECHNOLOGICAL ASPECTS

When assembling the prototype, a series of customizations have been done, starting with the receiver, which, in order to be capable of programming the coaster via radio required to have the XBee Explorer Dongle modified with the purpose of restarting the Arduino Fio remotely, a procedure which is essential to the operation. Therefore, a small jumper was welded between RTS and D3 pins, in such a way to bind the D3 status to the RTS, which, in turn, may be set to function as a reset pin of Arduino.

The power sensor chosen for the prototype has the biggest sensible area available in the market – 40 x 40mm – which allows the measurement of the pressure exerted by an object with an acceptable precision level. However, if the measured object is a bottle, one noticed that it required a material with a flat and hard surface applied to the sensor, preferably of the same size, to serve as a support, protection and ensure that there is pressure only in the useful area of the sensor.

A small modification has also been made in the sensor: machined sockets were placed in its outputs in order to facilitate its connection to the prototype, thus, avoiding the use of welding.

To feed the coaster, a 1000mAh Li-Po battery was used, which ensured to the prototype 12 hours of autonomy. However, depending on the alterations in the code – for example, by selecting the power economic mode of Arduino and turning on/off the XBee at every interaction with the receiver's XBee – we believe that the battery is capable of keeping the coaster in operation for about 120 hours.

Arduino Fio has undergone some adaptations. In the beginning, two rows of machined sockets were installed in the device, each on the existing sequence of pins, to facilitate the connection of wires, which resulted in a more practical and elegant solution.

Next, to avoid unnecessary welding and, thus, ensure the physical integrity and good appearance of Arduino Fio, a piece of standard board was used – attached with double-face adhesive tape to the surface of Arduino – to accommodate the other end of the wires, a resistor and a pair of machined sockets intended to work as a connector for the power sensor.

With respect to the prototype for the coaster, it is important to clarify that the power sensors are basically resistors changing the value of their resistance ( $\Omega$ ) according to the pressure exerted on it. Therefore, the simplest way to read a power sensor is to connect one of its outputs to a power source and the other to a resistor - in this case, of 10Kohm – connected to earth (GND). Next, the point between the resistor and the sensor should be connected to one of the analog inputs - in our case, an A0 input was chosen – of Arduino.

Since the working voltage of Arduino Fio is 3.3V, the analog voltage reader will vary from 0V (earth) and 3.3V, that is to say, the same voltage of the power source.

When the resistance of the power sensor falls, the total resistance of the sensor and the 10Kohm resistor also fall – of approximately 100Kohm to 10Kohm – which shows that, when the current passing through both resistors (sensor and resistor) increases, the voltage existing in the 10Kohm resistor increases accordingly.

In the end, the coaster prototype created, as shown in Figure 1, not only works with a reasonable faithfulness, but also occupies a small space (volume) – perhaps the smallest possible among the options with Arduino provided in the market – capable of assuming, except for its height, dimensions compatible with those of a common coaster, at a relatively reduced cost.

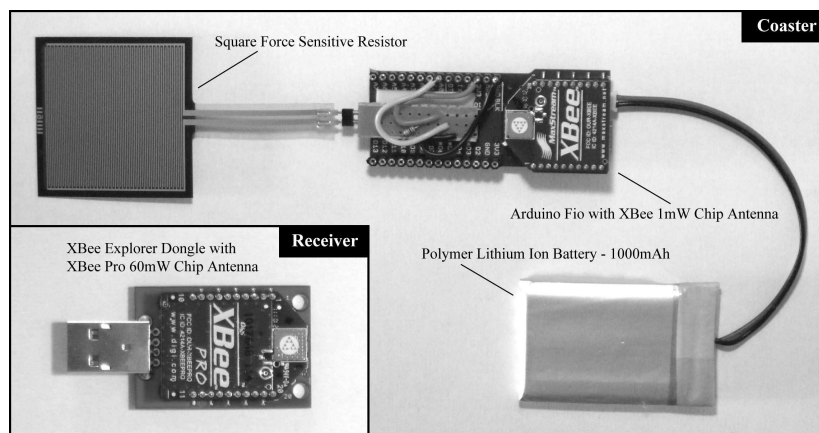


Figure 1. Prototype for the coaster and receiver

We created the code used for the test, which shows in the screen, at every 3 seconds, the amount of pressure recorded by the device. Specific messages are also established for each pressure stretches characterizing the container content level, without the need of making any calculation.

It is important to say that the messages created already suggest the use of that scheme in more than one coaster, and for each of them there is a unique identifier and an association to the table where it is located, in order to ensure the efficacy of the solution in a real situation. The result achieved may be noticed in Figure 2.

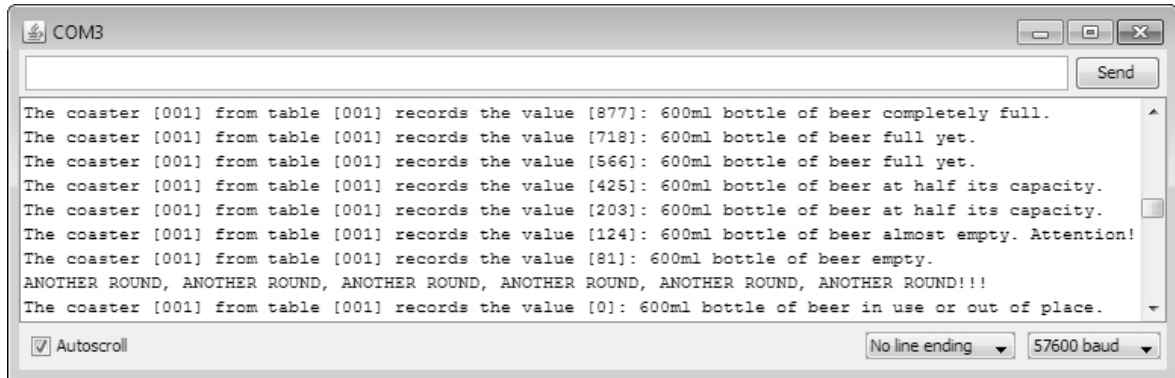


Figure 2. Graphic

## 5. CONSIDERATIONS ON THE DESIGN OF THE FINAL PRODUCT

About the coaster, the receiver and its interface with the user, some considerations should be taken. The coaster must be built in such a way to resist the contact with liquid, since its use supposes that beer would be spilled onto the equipment. Thus, the upper surface of the coaster and its side area should consist in a single part with a continuous surface without spaces. A similar solution should be adopted for the coaster base, establishing side parts to it. The material used could be high density polycarbonate or polyethylene.

Once the coaster enclosure is defined, the electronic parts located inside it will be fully protected due to the fact that no liquid spilled over the coaster will slide down by the sides of its upper part, being kept outside with no risk of infiltrating into the equipment, since the sides of its lower part will block such passage.

Another important aspect is to consider that the upper part cannot move horizontally in relation to the lower part, and the upper part should have a gap to fall down and pressure the power sensor. To solve this matter, the parts can be united by an internal spring system, which will provide sufficient resistance so that the sensor is capable of recording the zero and maximum pressure status, according to Figure 3.



Figure 3. Coaster enclosure

The coaster base should be made of some sort of material that can provide more attrition with the table, in order to avoid displacements. Thus, the application of a rubber film on the external surface is recommended.

Regarding the receiver, the data received must be changed into information for its final user: the waiter. Therefore, the final interface of the system should consider the quick diagnosis of the following variables:

- The priority to supply beer given by the waiting time of the next bottle: the longer the waiting time is, the more urgent the service should be. The presentation in a list format is a suitable visual solution;
- The reading of the interface by the waiter, which should be made quickly and, eventually, while he moves around the bar/restaurant. Therefore, the interface may be presented in a large screen that would serve not only to waiters, but also to customers;
- The position of the table in the bar/restaurant, to make it easy for the waiter to find it. A diagram representing the low plan of the site, in a schematic way, may be useful.



## 6. CONCLUSION, RELEVANCE AND DEVELOPMENTS

This project presents an evolution in the relationship between bars/restaurants and their customers. The result benefits both sides. The company's operation is optimized in such a way to generate more profits. The beer suppliers and producers also benefit from it just as the brands associated to the service (the bar/restaurant's beer brand or the product brand itself) incorporate values such as innovation, technology and pleasure.

More than a technological innovation, this proposal has a potential to become an innovation in the experience of entertainment among the public attending bars and restaurants. Using the information gathered by all tables may become, for instance, an argument to some kind of gratification for the tables completing a specific number of beers consumed, creating a game mechanics where the customer participates by simply being at the site. If the interface serving the staff is also visible to everybody in the bar/restaurant, the competition character will become even more evident when exhibiting the number of beers consumed by each table. In this case, this very interface may provide animations that emphasize the participants who are receiving more awards, in order to encourage customers to increase their consumption.

The innovative nature of the device entails the need of further studies, from its implementation in a real environment, in order to observe in more detail the consequences of its use and its social implications. From this point on, it is possible to promote changes in the proposed flow to offer a better consumer experience.

## REFERENCES

- Banzi, M., 2008. *Getting Started with Arduino*. O'Reilly Media, United States.
- Igoe, T., 2007. *Making Things Talk*. O'Reilly Media, United States.
- Noble, J., 2009. *Programming Interactivity*. O'Reilly Media, United States.
- Oxer, J., Blemings, H., 2009. *Practical Arduino: Cool Projects for Open Source Hardware*. Springer-Verlag, New York.

# CONCURRENT STUDY ON PERCEPTION AND BEHAVIOR OF DRIVER IN SIMULATED DRIVING

Hidetoshi Nakayasu<sup>1</sup>, Hirokazu Aoki<sup>2</sup>, Tetsuya Miyoshi<sup>3</sup> and Patrick Patterson<sup>4</sup>

<sup>1</sup>Konan University, 8-9-1 Okamoto Higashinada, Kobe 658-8501, JAPAN

<sup>2</sup>Hitachi Ltd., 890 Kashimada, Saiwai-ku, Kawasaki, 212-8587, JAPAN

<sup>3</sup>Toyohashi Sozo University, 20-1 Ushikawa, Matsushita Toyohashi, 440-8511, JAPAN

<sup>4</sup>Texas Tech University, 201 Industrial Engineering Building P.O. Box 43061 Lubbock, TX 79409

## ABSTRACT

A new approach was used to measure and analyze time histories of eye movements and vehicle response by use of an eye tracking system and driving simulator. The aim of this work is to determine the detection of hazard information in traffic situations to better forecast human driving behavior, such as visual attention. We studied the relation between histories of vision motion and vehicle behavior during turning scenarios in simulated driving.

## KEYWORDS

Perception of driver, Eye tracking system, Driving simulator, Human factors

## 1. INTRODUCTION

The driving system should be viewed as a man-machine system, made up not only of the movement of vehicles, but also of human perception, cognition, and the responses of the driver. In order to maintain a highly safe and reliable system of driving, the relationship between eye movements and driving behavior in conventional and unconventional situations should be studied. A survey on accidents shows that errors in visual perception are primary factors (Soliday, S. M., 1975, Miura, T., 1992, and Renge, K., 1998). Visual cognition requires the acquisition and judging of information. In order to develop a system which decreases the incidence of accidents, it is necessary to analyze the man-machine system with visual attention paid to the manner in which a driver operates a vehicle.

The experimental system proposed in this paper is able to realistically mimic the situation of driving and analyses human factors for awareness of hazard and risk. Another aim of this work is to offer the information regarding hazards in several different traffic situations, with a focus on human factors in order to forecast human behavior such as visual attention during driving task. In the experiments, the visual attention of driver was measured by the use of an eye-tracking system while participants operate a driving simulator. Eye movements and road scenes are analyzed from a moving driver's perspective. The relationship between eye movement and vehicle trajectory in Japanese road are also investigated.

## 2. EXPERIMENTAL SETUP

### 2.1 Driving Simulator and Eye Tracking System

A driving simulator with a 6-axis motion base system (Honda Motor DA-1102) was used to simulate driving an automatic transmission vehicle. Table 1 shows the specifications of the DA-1102. The simulator is controlled by a network of computers which also generate images from the driver's point of view, as



Figure 1. Schematic aspect of DS and Eye Tracking System.

well as those in the side-view and rear-view mirrors, car dynamics (such as steering and braking motions), and traffic scenarios. Ethernet-linked computers acquired a log of input and output variables (such as a steering, accelerator, brakes, the position of the car and speed). The instruments on the dashboard, steering wheel, gearshift, side brake, accelerator, and brake pedal were positioned in a manner similar to that in a real car, as shown in Figure 1.

This simulator enables one to drive a vehicle in scenario courses such as urban road, rural road, and highway and so on. Experimental studies were performed at an urban road scenario shown in Figure 2. This course consists of several traffic events; merging, passing and crosswalk with pedestrian, where dangerous situations were installed. Eye movements were recorded at 250 Hz with a spatial resolution of 0.022degree within a range of  $\pm 30$ degree in the horizontal axis and  $\pm 20$ degree in the vertical axis by an eye tracking system (SR Research Ltd. EyeLink II). Head movements were monitored, and the system automatically modified the data on eye position by using an eye camera. At the same time, the scene from the driver's perspective was recorded by a CCD camera equipped on a head mounted cap as shown in Figure 1.

## 2.2 Participants

Six male students (mean age = 23.3years old; SD = 0.84) participated in this experiment as drivers. The participants had standard Japanese automobile licenses and more than one year's worth of driving experience. All participants received more than 5 hours of training on how to use the driving simulator before the experiment. Prior to the experiment, the purpose of the study and the procedures were explained to the participants, and the informed consent of all participants had been obtained. All participants had normal or corrected-to-normal vision.

## 2.3 Experimental Procedures

Participants were seated in the driving simulator and drove the simulator in automatic transmission (AT) mode. After the participants practiced driving in the simulator for two or three runs on the urban road, an experimental session was conducted. In the experiment, participants were instructed to drive safely on the driving courses. On the urban road course the participants drove about 60 km/h.

In this study, horizontal and vertical eye movements of the participants were recorded from both eyes with 250Hz sampling rate in pupil and corneal reflection recording. At the same time, the physical data of vehicle such as current position, angle of wheel, degrees of depth of accelerator pedal and brake pedal can be recorded by 10 msec sampling rate.

Figure 2 shows the driving course of an urban road in Japan in which several traffic incidents are presented to the driver, as shown by the abbreviations ST, RT, LT, MG and STOP which, respectively, mean go straight, right turn, left turn, merge, and stop. The details of the incidents in Figure 2 are as follows. 1)

Table 1. Specifications of the DA-1102

Front view	Wide field (138 degree) screen projection
Rear view	3-mirror independent LCD display
CG	Redraw speed: 30 to 60 frame's
Mechanism	Six axis motion base system using G cylinders Control
Frame	Lightweight space frame structure with aluminum extended mechanism
Body	Rear open structure fixed with FRP mold
Operation system	Steering device with reactive force control, accelerator, clutch, brake
Mission	AT/MT switch mechanism
CG computer	Cuamum3D Alchemy (1.5 million polygons)
Control system	6-axis servo amp
Dimensions	2,440mm(D) × 2,280mm(W) × 1,855mm(H)

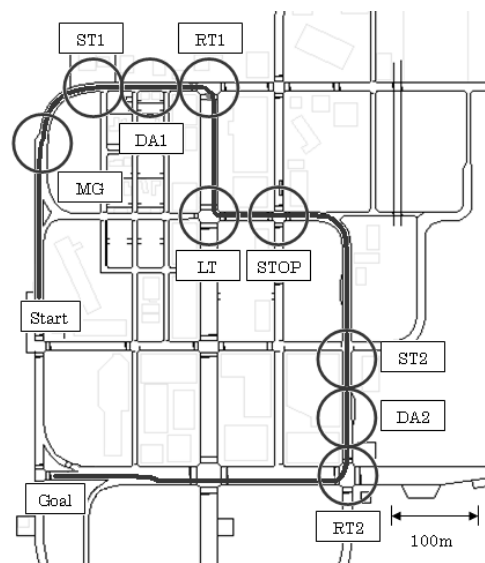


Figure 2. Driving course with implemented traffic events

The course featured a number of dangerous incidents such as car suddenly swerves as a child jumps out into the car path; 2) The course featured fine weather as clear as day.

These data were analyzed on the basis of velocity and position for each traffic incident to evaluate the mean distance of eye movement, fixation frequency and duration. This paper defines a saccade of eye movement occurring as when an eye moves at a velocity of more than 90deg/sec. An eye moves after the occurrence of a saccade. This paper also defines a fixation of the eye as when an eye does not move for more than 48msec after an eye movement and a saccade. Therefore, the fixation time refers to the duration time from the saccade after an eye movement until the saccade before an eye movement.

### 3. RESULTS AND DISCUSSIONS

#### 3.1 Vehicle Trajectories and Eye Movements

In Figure 3, the relationship of eye movement data such as the axis of the driver's eye (X-eye, Y-eye) and driver behavior (Steering angle, Throttle, Brake) such as steering angle, degree of throttle and brake are illustrated on the same axis for comparison. Therefore, different units of measurement were used in the vertical axis of Figure 3. In the figure, X-eye expresses the horizontal eye movement as the eye angle (degree) unit. Y-eye refers to the vertical eye movement of the eye angle (degree) unit. This figure shows the behavior of drivers at the right turn (RT1) incident. Vehicle trajectory during the right turn (RT1) incident is also drawn in Figure 4 with an eye fixation time stamp, where one of the video clips shows the scanning behavior of a driver reacting to the scene as shown in Figure 5.

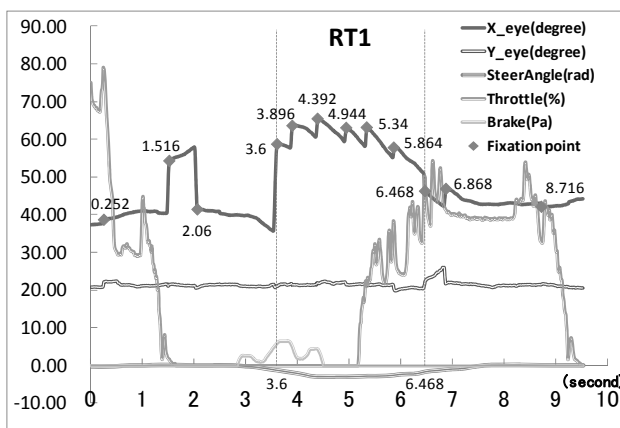


Figure 3. Eye movement and log data of DS at RT1

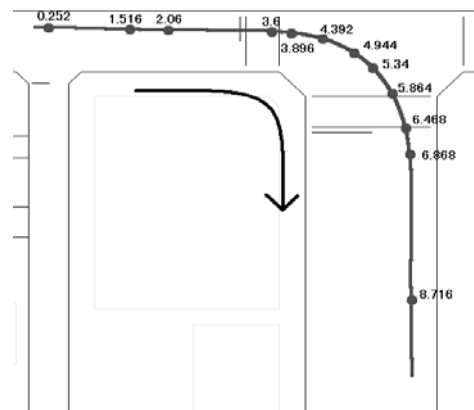


Figure 4. Vehicle trajectory at RT1

It is found in Figure 3, from 0sec to 3.60sec that the driver became aware of a vehicle coming towards him or a person crossing. The driver in this section confirmed that it was safe before making a right turn. We defined this as the introduction stage of turning. In section Figure 3 from 3.60sec to 6.468sec, the value of X-eye increases. Generally, it is known that the human eye repeats a saccade (rapid eye movement) and the stop (T Brown, I.D. and Groger J.A., 1988). It is shown that saccade and "smooth pursuit eye movement (The Vision Society of Japan Eds, 2001) are repeated rapidly in the turning behavior of driver. For the driver, the forward appearing scenery begins to drift. Therefore, the driver's eye movement must also drift. This is referred to as smooth pursuit eye movement. Figure 5 features two video clips which show the observation position of drivers at  $t = 5.34(\text{sec})$  and  $t = 5.80(\text{sec})$ . The filled circle in the figure expresses the fixation point of the driver. The eyes engage in multiple small smooth pursuits' eye movements. The driver's eye movements demonstrate that he/she can only see in front of the vehicle. This behavior is defined in this paper as the practice confirmation stage of turning. In section Figure 3 from 6.468sec to 9.52sec, the driver acquires information about the road after turning. This behavior can be referred to as the completion stage of turning.

The above suggests that there are three phases in the turning behavior of a driver: the introduction,

practice confirmation and completion stages. At the introduction stage, the driver performs perception and the cognition. At the practice confirmation stage, the driver judges, practice and confirms his/her actions. At the completion stage, the driver confirms the completion of the operation and then shifts to the next target, once again moving back to the introductory stage.

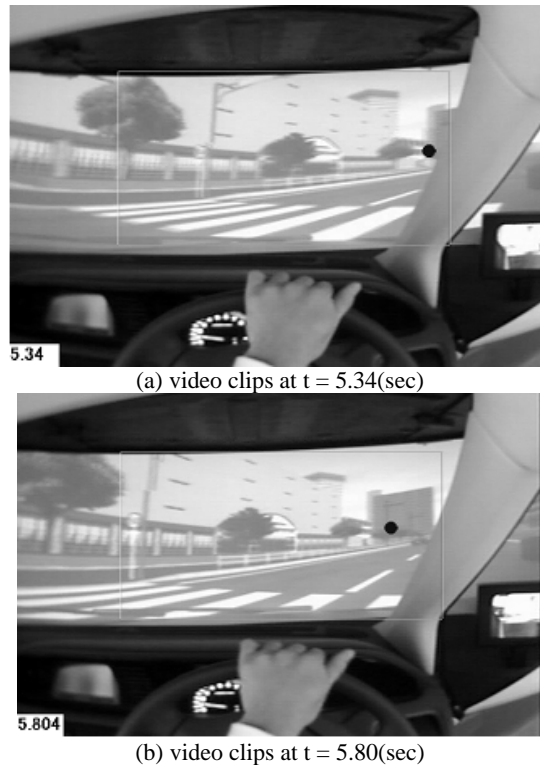


Figure 5. Video clips of driver's view

### 3.2 Eye Movement and Duration Time of Fixation

Figure 6 shows the mean value of average eye movement distance for the six participants in each traffic incident, where the vertical line designates standard errors (SE). The average eye movement distance refers to the mean values of the average number of fixations and the duration of fixations for the six participants, as illustrated in Figures 7 and 8. The average frequency of fixations refers specifically to the average number of fixations that occurred during selected traffic incidents. On the other hand, the average duration of fixation refers to the average time required for each of the selected traffic incidents.

In Figure 6, the vertical line shows the standard error (SE). From Figure 6 it is seen that the mean distance of eye movement at the merging incidents (MG) is much larger than for other traffic incidents. It is seen that in the MG incident that the driver sometimes watched a side mirror and a forward direction in merging point of road. However, the same behavior was not found in the straight road (ST1). As a typical example,

From Figures 7 and 8, it is noted that for most traffic events that when the fixation frequency is small the duration of fixation become large. For example, the fixation frequency at STOP is smaller than that at other traffic incidents, while the fixation duration at STOP is longer than that at other traffic incidents. On the other hand, the fixation frequency at MG and ST1 is larger than for other traffic incidents, while the duration time at MG and ST1 is smaller than for other traffic incident. These tendencies show that when there is little fixation frequency, the driver must collect a great deal of information at once by focusing their attention.

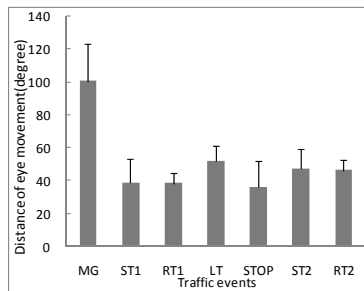


Figure 6. Distance of eye movement.

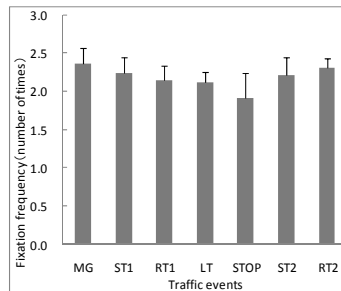


Figure 7. Fixation frequency.

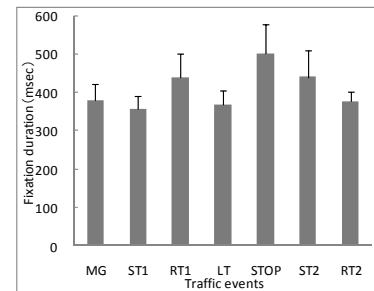


Figure 8. Duration Time of Fixation

#### 4. CONCLUDING REMARKS

A study was conducted to provide information about search behavior, useful for understanding its role in accident prevention. Eye movement and the trajectory of vehicles were analyzed in a variety of simulated driving scenarios. We found the following:

1. Three phases were found in the turning behavior of a driver: the introduction, practice confirmation and completion stages. At the introduction stage, the driver performs perception and the cognition. At the practice confirmation stage, the driver judges, practice and confirms his/her actions. It is shown that saccade and “smooth pursuit eye movement are repeated at practice confirmation stage in turning behavior of driver. At the completion stage, the driver confirms the completion of the operation and then shifts to the next target, once again moving back to the introductory stage.

2. We found, for most traffic events, that when the fixation frequency was small the duration of fixation became large. Conversely, when there were smaller fixation frequencies, the driver must collect a great deal of information at once, focusing attention resources within a limited field of view.

The relationship between eye movement and vehicle trajectory in conventional and unconventional driving situations was studied. Future studies will investigate the relationship of these situational search patterns to accidents and driver error.

#### REFERENCES

- Brown, I.D. and Groger J.A., 1988. Risk perception and decision taking during the transition between novice and experienced driver status, In *Ergonomics*, Vol.31, pp.585-597.
- Miura, T., 1992. Visual Search in Intersections – An Underlying Mechanism -, In *IATSS RESEARCH*, Vol.16, pp.42-49.
- Renge, K., 1998. Drivers’ hazard and risk perception, confidence in safe driving, and choice of speed, In *IATSS RESEARCH*, VOL.22, pp.103-110.
- Soliday, S. M., 1975. Development and preliminary testing of a driving hazard questionnaire, In *Perceptual & Motor Skills*, VOL.41, pp.763-770
- Edited by The Vision Society of Japan, 2001. Handbook of visual information process, pp.91–92., pp.393–398. (in Japanese)

# OPTIMAL POSITIONS OF ADVERTISEMENTS ON NEWS WEBSITES FOCUSING ON THREE CONFLICTING OBJECTIVES

Noriyuki Muramatsu, Keiki Takadama and Maki Sakamoto  
*Department of Informatics and Engineering*  
*The University of Electro-Communications*

## ABSTRACT

This paper proposes optimal positions of advertisements on news websites. In this paper we take into account the advertising effectiveness measured by the attention degree, the impression degree and the usability degree of news websites at the same time. Since previous studies show that the attention degree and the impression degree of advertisements are in the trade-off relation, the goal of this study is to show the effective layouts of the advertisements on news websites optimizing three conflicting objectives. We pursue this goal by conducting experiments in which the participants view a variety of page layouts of the news websites. The results of experiments on the attention degree, the impression degree and the usability degree were analyzed from the viewpoint of the multi-objective optimization. As a result of the analyses, we got the six optimal layouts of the advertisements.

## KEYWORDS

Advertisement, news websites, optimization, attention, impression, usability

## 1. INTRODUCTION

Advertisements (hereafter, ads) in news websites today are very important because they are watched by the majority of net users. (Edelman 2007) claims the necessity of specific strategies for how to best use websites as an advertising medium. The ads in websites should attract users' attention (Janoschka, 2004). When the ads in websites try to appeal to users, however, they compete with other elements and contents like articles, headlines, illustrations, etc., that are also placed on the websites. From these viewpoints, the eyetrack III research released by the pointer institute observed where and how the ads are effective. The result suggests that people often ignore ads, but good ads placement and appropriate format for the ads can improve the users' attention degree of ads. However, it is also important to consider what kind of impressions such as emotions are evoked by the ads as a result of getting attention. (Muraoka et al. 2010) and (Sakamoto and Takadama to appear) assumed that the attention degree and impression degree of ads in news websites are in the trade-off relation, and analyzed the attention degree and the impression degree of ads from the viewpoint of the multi-objective optimization. Previous studies of ads have pursued the effective web layouts for ads. In the news websites, however, the goal of users is finding and reading their necessary information. (Sakamaki et al. 2009) point out the importance of the usability of news websites. The usability degree of the news websites decreases when ads are inserted in the high attention placement. Therefore, in this study we explore the effective placements of ads in news websites, which give the high attention degree, the high impression degree, and the high usability degree at the same time. We pursue this goal by conducting experiments in which participants view variety of pages layouts of news websites. These results are analyzed from the viewpoint of the multi-objective optimization.

This paper is organized as follows; the next section and Section 3 describes the procedure of the subject experiment and its results. Section 4 analyzes the experimental results from the viewpoint of the multi-objective optimization, and conclusion and suggestions for future work are given in Section 5.

## 2. PSYCHOLOGICAL EXPERIMENT

In this study, we conducted psychological experiments to explore effective placements of ads in new websites. The participants of the experiment watch various types of newssite samples in which ads are positioned in various layouts. Through the experiments, we measured (1) the eye fixations on ads, (2) the impression of ads related to the contents of the negative news articles and (3) the usability of news websites.

We used six kinds of ad categories with the high insertion frequency for three months: (1) the service; (2) the finance; (3) the real estate; (4) the information and communication; (5) the cosmetics; and (6) the leisure. The layouts employed in the experiments were the ten patterns as shown in Figure 1, where the bold square indicates the advertisement.

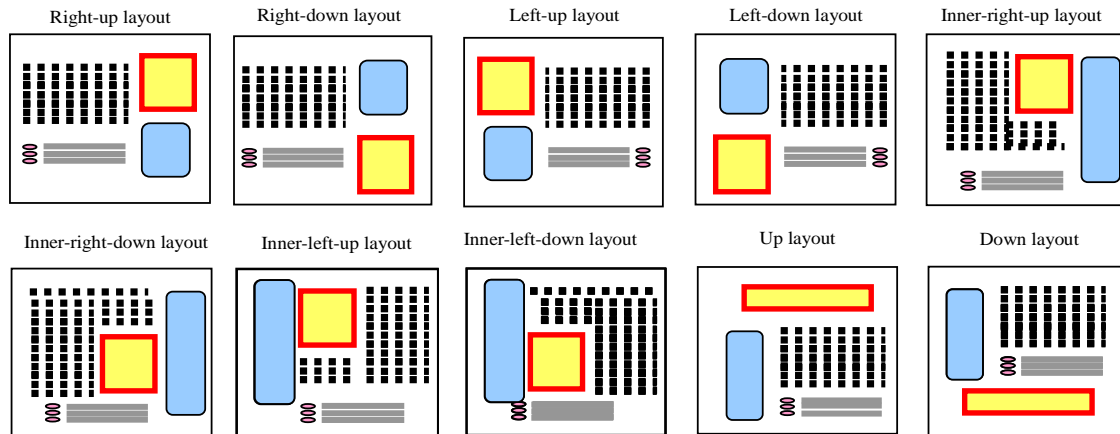


Figure 1. Ten patterns of ad layouts used for the experiments

We conducted the following three types of psychological experiment in order to explore effective placements of ads in new website. In each experiment, 20 participants watched randomly 12 newssite samples.

□ **Experiment 1:** We measured eye movements of the participants using an eye tracking equipment. The newssite samples were selected randomly and there was no time restriction for users in watching the newssite samples. In this experiment, the following three types of eye fixation data were measured: (1) the number of time of the eye fixation; (2) the eye fixation duration; and (3) the rate of the participants who fixed their eyes on the ad at least once.

□ **Experiment 2:** We measured the impression of ads in news websites showing the negative news articles. We conducted the questionnaire by asking participants to evaluate the impression of ads after watching the news website including the ads. In the questionnaire, we employed the seven-points SD scales from extremely negative (-3) to extremely positive (+3).

□ **Experiment 3:** We measured the usability of news websites showing negative articles. We conducted the questionnaire in which participants were asked to evaluate the usability of the news websites.

## 3. RESULTS OF EXPERIMENTS

Table 1 shows the result of the experiment 1. In this table, line indicates the three evaluation criteria on the eye fixation on the ads, while the column indicates the ten types of web layouts. The layouts that are high in value are the inner-right-up layout and inner-left-down layout. This result indicates that eyes move across ads when ads are located at the positions near to the news articles users read.



Table 1. The mean values of number of fixation, fixation duration, and the rate of participants who watch ads

	Right-up	Right-down	Left-up	Left-down	Inner-right-up	Inner-right-down	Inner-left-up	Inner-left-down	Up	Down
number of eye fixations	0	1.13	1.32	1.21	3.26	2.45	1.33	3.55	0	0.9
eye fixation duration	0	0.13	0.2	0.16	0.6	0.3	0.21	0.48	0	0.15
rate of participants who watched ads	0	0.35	0.52	0.46	0.6	0.73	0.62	0.55	0	0.4

Table 2 shows the result of the experiment 2. In this table, the line shows the two types of SD scales. The up layout is high in value and the inner-right-down layout is low in value. This result is conflicting to the result of experiment 1.

Table 2. The mean values of ads impression for 2 SD scales

	Right-up	Right-down	Left-up	Left-down	Inner-right-up	Inner-right-down	Inner-left-up	Inner-left-down	Up	Down
ads impression	-0.33	-0.58	0.04	-1.04	-0.67	-0.71	-0.54	-0.79	-0.83	-0.83
purchase intent	-0.33	-0.46	-0.46	-0.71	-0.46	-1.13	-1.04	-0.96	-0.13	-0.71

Table 3 shows the result of the experiment 3. In this table, the line shows the two types of SD scales. The layouts that are high in value are the down layout and the right-down layout, while the inner-left-up layout is low in value. Therefore, this result indicates that the attention degree and the usability degree are conflicting.

Table 3. The mean values of newssite usability for 2 SD scales

	Right-up	Right-down	Left-up	Left-down	Inner-right-up	Inner-right-down	Inner-left-up	Inner-left-down	Up	Down
easiness in reading articles	1.08	1.63	1.38	1.17	0.42	0.38	-0.13	0.63	1.3	1.42
simplicity of the newssite layout	1.29	1.3	1.17	1.71	0.8	0.33	0.58	0.8	1.21	1.5

## 4. MULTI OBJECTIVE OPTIMIZATION

### 4.1 Setting up Three Objectives

The strong positive correlation coefficient among the three types of eye fixation data was observed ( $r=0.959, p < .05$ ). Based on this result, we summarized the three types of eye fixation data as an “attention degree”. The relatively strong positive correlation coefficient among the two types of impression data ( $r = .555, p < .05$ ) and the strong positive correlation coefficient among the two types of usability data ( $r = .817, p < .05$ ) were observed. Therefore, we summarized the two types of impression data as an “impression degree” and the two types of usability data as a “usability degree” respectively. Here, we define the “attention degree” as the degree to which the ads attract the eyes of users/consumers, the “impression degree” as the degree to which the ads are or are not influenced by the negative news articles and the “usability degree” as the degree to which news articles are or are not easy for users to read them.

We calculated the correlation coefficients among the three scales of the attention degree, impression degree and usability degree. As a result, the attention degree and the impression show the strong negative correlation. The attention degree and the usability degree show the relatively strong negative correlation. Furthermore, the impression degree and the usability degree show the equilateral weak correlation.

## 4.2 Analysis of Three Objective Functions

We explored the Pareto front solutions which are better solutions than the others in the three objective functions, namely the attention degree, the impression degree and the usability degree. We call “x dominate y” ( $f(x) \geq f(y)$ ) when two solutions satisfies the following two equations, where  $f_i(x)$  indicates the solution x on i-th objective function:  $\forall_i \in \{1, 2, \dots, m\}: f_i(x) \geq f_i(y) \wedge \exists_i \in \{1, 2, \dots, m\}: f_i(x) > f_i(y)$ . Here the m is 3 because we have three objective functions.

From this viewpoint, we explored the Pareto front solutions, whose values are shown in Table 4. The shaded scores are the highest score for each objective function.

Table 4. The score of the Pareto front solutions

	left-up	up	under	Right-down	Inner-right-up	Inner-left-down
attention degree	0.47	0.00	0.35	0.34	0.91	0.85
impression degree	-0.21	-0.17	-0.77	-0.52	-0.56	-0.88
usability degree	1.27	1.25	1.46	1.46	0.60	0.71

The characteristics of these six types of web layouts are summarized as follows:

- Left-up layout

This layout provides a relatively high score of the attention degree, impression degree and usability degree. The ads located at the left-up layout gains high attention degree. We suppose that this is because users start to read from the left side in Japanese in the same way as English and read the article’s title first. The ads in this layout also gained high impression degree. We suppose that in this layout the negative news articles do not influence on the impression of ads because users are not reading the negative news articles when they see ads. This layout does not disturb users in reading the news articles and keeps the usability degree high.

- Up layout

This layout has provides the high impression degree and the relatively high usability degree. However, the attention degree is low because the ad located at the up position is not frequently watched by users. This result indicates that the ad in the Up position is not influenced by the negative contents of articles because it is not seen by users.

- Down layout

This layout provides the low impression degree, the relatively high attention degree and the high usability degree. We believe that the impression degree is low because users evaluate the impression of ads after reading the negative news articles. However, this layout does not disturb users in reading the news articles and keeps the usability degree high.

- Right-down layout:

This layout has the low attention degree, the relatively high impression degree and the high usability degree. Ads located at the right-down position are not frequently watched by users. This indicates that ads in this position are not influenced by the negative contents of articles. This layout also does not disturb users in reading the news articles and keeps the usability degree high.

- Inner-right-up layout

This layout has the relatively low impression degree and usability degree. On the other hand, the attention degree is high. This suggests that ads located at the inner-right-up position are frequently watched by users. Therefore, the impression of ads in the inner-right-up position is easily influenced by the negative news articles because the ads in this layout tend to be watched frequently during reading the negative news articles.

- Inner-left-down layout

This layout has the low impression degree and usability degree, while the attention degree is high. This suggests that ads located at the inner-left-down position are frequently watched by users and the impression of ads in the inner-left-down position is easily influenced by the negative news articles.

What should be noted that the right-up layout, which is the current standard layout, do not become the Pareto front solutions. This indicates that the ads in the current standard layout are directly affected by the negative news articles and the standard layout is not appropriate when negative news articles are posted in the website.

## 5. CONCLUSION

Our research explored the optimal positions of ads on news websites showing the negative news article related to the content of ads. We took into account the advertising effectiveness measured by attention degree and impression degree and the usability of news websites at the same time. Since previous studies show that the attention degree and the impression degree of ads are in the trade-off relation, the goal of our study was to show the effective layouts of the ads on news websites optimizing three conflicting objectives. We pursued this goal by conducting experiments in which the participants view a variety of page layouts of the news websites. The results of experiments on the attention degree, the impression degree and the usability degree were analyzed from the viewpoint of the multi-objective optimization. As a result of the analyses, we found the six optimal layouts of the ads: (1) left-up layout; (2) up layout; (3) down layout; (4) right-down layout; (5) inner-right-up layout; and (6) inner-left-down layout. Our results indicate that the ads located near to the news articles gain high scores in the attention degree, while they have low scores in the impression degree and the usability degree. Our results also suggest that the right-up layout, which is the current standard layout, is not a good layout when ads are inserted into the negative news articles.

The results of this paper are based on ten types of the web layouts. Further careful qualifications and justifications, such as analysis of other types of the web layouts, are needed to generalize our results from the viewpoint of multi-objective optimization. (Oliver et al. 2002) proposes the use of an interactive genetic algorithm which generates styles or layouts of web sites, though their study does not deal with ads. Therefore, another important direction to be pursued in the near future is to explore optimal positions of ads on web sites by utilizing the Pareto front solutions through Genetic Algorithm or Interactive Genetic Algorithm.

## REFERENCES

- Oliver, A., Monmarche, N., and Venturini, G., 2002. Interactive Design of Web Sites with a Genetic Algorithm. *IADIS International Conference WWW/internet 2002*, pp355-362.
- Edelman, David C., 2007. From the Periphery to the Core: As Online Strategy Becomes Overall Strategy, Marketing Organizations and Agencies Will Never Be the Same. *Journal of Advertising Research*, Vol.2, No.47, pp130-134.
- Janoschka, A., 2004. *Web Advertising*. John Benjamins Publishing Company.
- Muraoka, K., Sakamoto, M. and Takadama, K., 2010. The Layout Optimization of Newssite-Inserted Advertisements Using Two Conflicting Objectives. *Proceedings of the IADIS International Conference Interfaces and Human Computer Interaction*, pp147-154.
- Outing, S., and Ruel, L. 2004. Eyetrack III – What You Most Need to know-. <http://poynterextra.org/eyetrack2004/main.htm>. USA.
- Sakamaki, R., Someya, S. and Okamoto, K., 2009. An Analysis of Relationship between Evaluation Structure and Memorized Information Quantity on the Web-site. *Japanese Society for the Science of Design*, Vol.55, No.6, pp59-66.
- Sakamoto, M. and Takadama, K. (to appear) The layout Optimization of Newssite-inserted Advertisements Using Two Conflicting Advertisements. *International Journal of Computer Information Systems and Industrial Management Applications*.

# IMPROVEMENT OF DETECTION PROBABILITY OF INSPECTION PANEL BY MULTIPLE MODARITIES

Hidetoshi Nakayasu<sup>1</sup>, Masao Nakagawa<sup>2</sup>, Tetsuya Miyoshi<sup>3</sup> and Patrick Patterson<sup>4</sup>

<sup>1</sup>*Konan University*

*8-9-1 Okamoto Higashinada, Kobe 658-8501, JAPAN\**

<sup>2</sup>*Shiga University*

*1-1-1 Bamba, Hikone, Shiga 522-8522, JAPAN*

<sup>3</sup>*Toyohashi Sozo University*

*20-1 Ushikawa, Matsushita Toyohashi, 440-8511, JAPAN*

<sup>4</sup>*Texas Tech University*

*201 Industrial Engineering Building P.O. Box 43061 Lubbock, TX 79409*

## ABSTRACT

A procedure for measurement of psychometric function of a single modality and joint probability of dual modality in the inspection image data is proposed. The validity of the proposal method was also verified by the example of the experiment of single and dual modalities. The differences of the detection probabilities were illustrated by the features in single modality and multiple modalities.

## KEYWORDS

Multiple Modalities, Detection Probability, Lower bound by multiple modalities, Correlation

## 1. INTRODUCTION

Several kinds of image processing method have been applying to the automated visual inspection system for defects on the product surface inspection (Chin, et. al., 1982), since the precision and speed will be degraded by fatigue of the inspector. In order to meet these problems, some research works have been tried for the standardization of an operation time of visual inspection (Morawski, 1992, and Drury, 1972).

In the practical visual inspection, a panel must decide the judgement of the quality by using one's sensitivities for multiple modalities whether the quality of object is good or not. In order to improve the detection probability of inspection panel by multiple modalities, a new visual sensory test is proposed under the situation of various combinations of mixed modalities.

## 2. THEORETICAL APPROACH OF DETECTION PROBABILITY

### 2.1 Detection Probability with Multiple Modalities and Its Lower Bound

The probability obtained by single modality is formulated by psychometric curve(Hayashi H. et al., 2007) that means the probability that one can feel the strength of stimuli of target object less than specified strength of stimuli such as threshold  $x^0$  in single modality as shown in Figure 1. This probability is defined by

$$P_i = P(E_i) \quad (1)$$

where  $P(\bullet)$  is an occurrence probability of probability event  $\bullet$ .  $E_i$  is the probability event that a panel judges that the strength of the stimuli of the target sample is less than that of the threshold  $x_i^0$  such as a standard sample for  $i$ -th modality among multiple modalities.

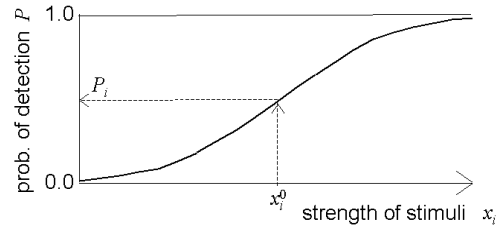


Figure 1. Psychometric curve

On the other hand, now suppose that the detection probability  $P$  under the situation of multiple modalities are defined by

$$P = P(E_1 \cup E_2 \cdots \cup E_n) \quad (2)$$

where  $P_1 \leq P_2 \leq \cdots \leq P_n$ . This equation means the probability that the panel can detect the strength of stimuli less than specified strength by using one of modalities at least.

In this condition of  $P_1 \leq P_2 \leq \cdots \leq P_n$ , the lower bound probability  $P_L$  under multiple modalities can be estimated by the equations. (Nakayasu et al., 2007).

$$P_L = P_1 + (P_2 - P_{12}) + \sum_{k=3}^n \max \left\{ P_k - \sum_{l=1}^{k-1} P_{lk}, 0 \right\} \quad (3)$$

Eq.(3) is the good estimator of probability  $P$  with multiple modalities.

In Eq.(3),  $P_{lk}$  is the joint probability of the events  $E_l$  and  $E_k$

$$P_{lk} = P(E_l \cap E_k) \quad (4)$$

This probability is derived by the integral

$$P_{kj} = \Phi(y_k^0, y_l^0) = \int_{-\infty}^{y_k^0} \int_{-\infty}^{y_l^0} \phi(y_k, y_l; \rho_{kl}) dy_l dy_k \quad (5)$$

In the equation,  $\Phi(y_k^0, y_l^0)$  is the cumulative probability of standardized two-dimensional normal distribution.  $\phi(y_k, y_l; \rho_{kl})$  is the joint probability density function of two-dimensional standard normal distribution such as:

$$\phi(y_k, y_l; \rho_{kl}) = \frac{1}{2\pi\sigma_k\sigma_l\sqrt{1-\rho_{kl}^2}} \exp \left\{ -\frac{1}{2(1-\rho_{kl}^2)} (y_k^2 - 2\rho_{kl}y_k y_l + y_l^2) \right\} \quad (6)$$

$$y_k = \frac{x_k - \mu_k}{\sigma_k}, \quad y_l = \frac{x_l - \mu_l}{\sigma_l}, \quad y_k^0 = \frac{x_k^0 - \mu_k}{\sigma_k}, \quad y_l^0 = \frac{x_l^0 - \mu_l}{\sigma_l}, \quad \rho_{kl} = \frac{\sigma_{kl}}{\sigma_k \cdot \sigma_l} \quad (7)$$

$$\mu_{(.)} = \int_{-\infty}^{\infty} x_{(.)} \cdot f(x_{(.)}) dx_{(.)}, \quad \sigma_{(.)}^2 = \int_{-\infty}^{\infty} (x_{(.)} - \mu_{(.)})^2 f(x_{(.)}) dx_{(.)}, \quad \sigma_{kl} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_k - \mu_k)(y_l - \mu_l) f(x_k, x_l) dx_k dx_l \quad (8)$$

$$f(x_k) = \frac{1}{\sqrt{2\pi}\sigma_k} \exp \left\{ -\frac{(x_k - \mu_k)^2}{2\sigma_k^2} \right\} \quad (9)$$

## 2.2 Psychometric Function by Single Modality

The psychometric curve is defined by the strength of stimuli as shown in Figure 1.

$$P = f(x), \quad P = \Phi(y) = \int_{-\infty}^{y_i} \phi(y) dy, \quad y = \frac{x - \mu_x}{\sigma_x}, \quad \phi(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \quad (10)$$

$\Phi(\bullet)$  and  $\phi(\bullet)$  are cumulative distribution function and probability density function of the

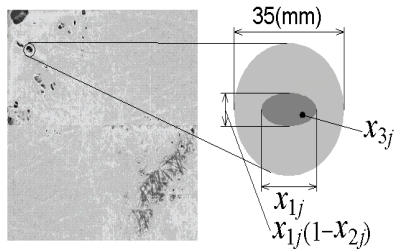


Figure 2. Artificial replica for three kinds of modalities created from FRP inspection image data

Table 1. Statistical values by image processing

	area size (mm <sup>2</sup> )	aspect ratio	color density
mean	$\mu_A = 5.855$	$\mu_R = 0.361$	$\mu_V = 0.547$
S.D.	$\sigma_A = 19.113$	$\sigma_R = 0.197$	$\sigma_V = 0.422$

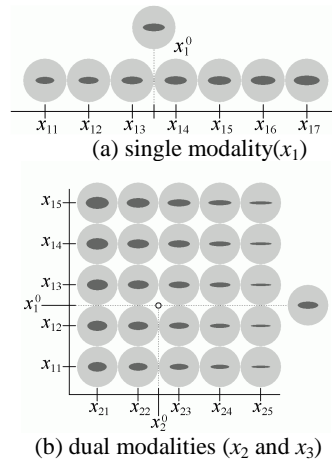


Figure 3. Examples of degree of stimuli for single and dual modalities sensory test

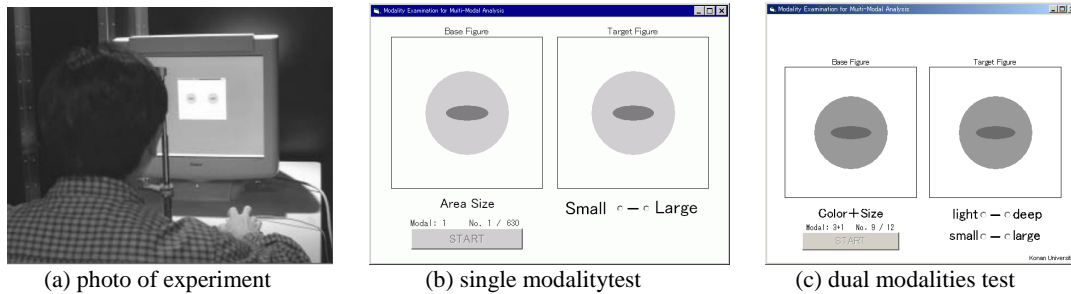


Figure 4. Schematic aspect of sensory test

standardized normal distribution.  $\mu_x$  and  $\sigma_x$  are a mean and standard deviation.

### 3. EXPERIMENTAL APPROACH OF DETECTION PROBABILITY

In this study, the task and human perceptual ability to FRP (fiber reinforced plastics) product surface inspection image data as shown in Figure 2 is taken into account in order to evaluate the human sensitivities on each modality and the correlation among modalities. Three modalities such as size, aspect ratio, and grayscale of an inspection object were manipulated by the artificial values produced for sensory test by using the values in Table 1 obtained practical inspection data.

As to the amount of the feature, the first modality  $x_1$  means the size whose value is dependent on the length of the long axis of an ellipse. The second modality  $x_2$  is the shape defined by aspect ratio of the length of the long axis to short axis of an ellipse. Finally, the third modality  $x_3$  is quantified by the degree of grayscale. In this way, the strength of the stimuli for each modality  $x_{ij}$  ( $i=1,2,3, j=1,2,\dots$ ) was determined as a set for the paired comparison. In Figure 3, the visual aspect of degree of strength of single modality and mixed modalities is drawn, but the values of modalities used for Figure 3 are the samples for typical replica for the explanation. The details of data treatment and numeric conversion of stimulus strength of these three modalities can be referred by Nakayasu et al. (Nakayasu et al., 2007).

31 participants of 20-26 years old joined to the experiments of single and dual modalities sensory tests with informed consent. In the experiments, the head of participant is fixed shown in Figure 4(a) and the stimulus of sensory test appeared in the display located with the distance of 60cm from the participant's eyes. There are two kinds of stimulus used in sensory test as shown in Figure 4 (b) and (c) that was generated by the PC with the programming code developed by Visual Studio 6.0 (Visual Basic language) on Windows OS.

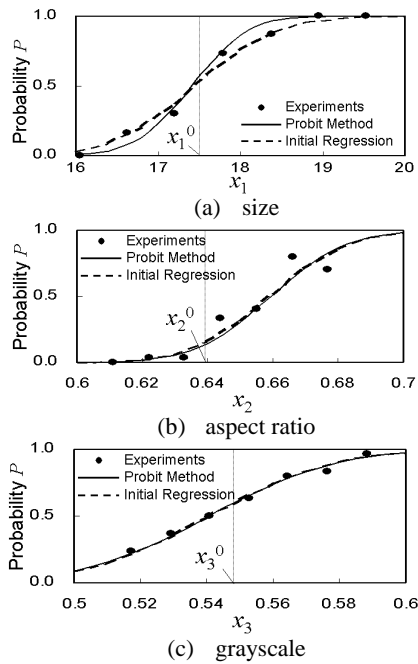


Figure 5. Psychometric curves by single modality

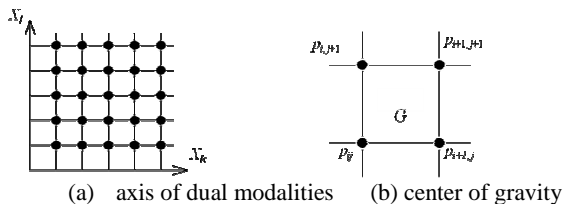


Figure 6. Discrete observation probability and definition of axis

## 4. RESULTS & DISCUSSIONS

### 4.1 Prob. by Single Modalities

It is seen from Figure 5 that the psychometric curves obtained by single modality is almost fitting in initial regression curves in the case of  $x_2$  and  $x_3$  while the curve is not fitting in initial regression curve in the case of  $x_1$ . The reason for the discrepancy in the modality  $x_1$  is due to the experimental data of modality  $x_1$ . From the obtained psychometric curves, it will be forecasted that the sensitivities for the size and aspect ratio are slightly higher than that of grayscale in Figure 5, since the gradient of the curves are larger than that of grayscale.

### 4.2 Detection Probability by Multiple Modalities

From the data measured by paired comparison test for dual modalities, the correlation matrices among  $x_1, x_2$  and  $x_3$  are listed in Figure 7 for typical subjects from 10 subjects. It is seen from the table that there are three kinds of typical patterns such as negative((a)subject A), positive((b)subject B) and less((c)subject C) correlations among  $x_1, x_2$  and  $x_3$ .

Calculation results are illustrated in Figure 7 (a),(b) and (c) for typical three subjects that are corresponding to those listed in Table 1. In all figures, the detection probabilities using multiple modalities are larger than that using single attribute. For the evaluation of lower bounds using multiple modalities, it is useful to estimate the effect to detection probability of three kinds of modalities, since it is seemed that the

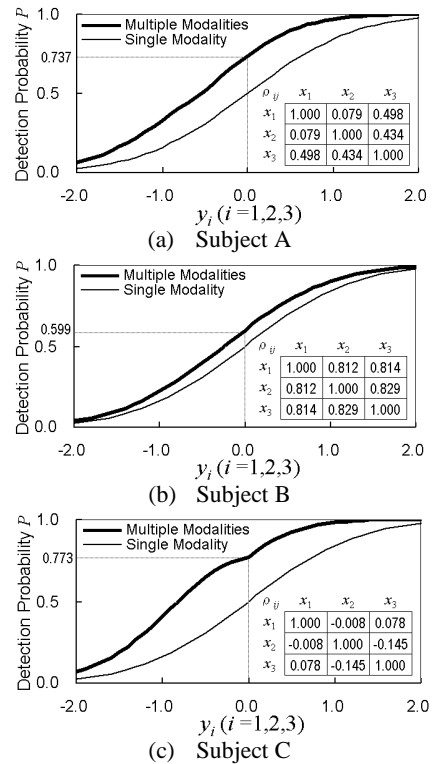


Figure 7. Psychometric curves by multiple modalities

upper bound of the probability for multiple modalities is overestimated. Therefore it is recommended that the evaluation of probability must be considered by the lower bound of probability for multiple modalities. It is also interesting that the lower bound of probability for multiple modalities with positive correlation yields a little rate of increase rather than the case with negative and less correlation, though the probability increases at the case of dual and multiple modalities.

From the results in the experiments, the probabilities for dual and multiple modalities are larger than those using single modality. This tendency is same in all experimental conditions. Increasing amount of detection probability, however, is not same in different ways with the variation of correlation coefficient. In fact, the variation is seen from the correlation coefficient as shown in Table 3. Table 4 shows that the improved ratio of detection probabilities is strongly dependent on P<sub>12</sub>, P<sub>23</sub>, P<sub>31</sub> using dual modalities: respectively rise approximately 21%, 26% and 24% in the case using dual modalities experiments. That is to say improving ratio of detection probabilities have a pronounced tendency to depend on correlation coefficient of a combination of dual modalities.

Table 3. Correlation of coefficient

Subject	$\rho_{12}$	$\rho_{23}$	$\rho_{31}$
A	-0.406	-0.426	0.397
B	0.072	0.173	0.1
C	0.545	-0.522	-0.718
D	0.423	-0.042	0.528
E	0.532	-0.249	-0.315
F	0.573	-0.18	0
G	0.177	-0.373	-0.104
H	0.342	0.224	-0.272
I	0.528	0.548	0.048
J	-0.1	0.031	-0.02
K	-0.038	-0.095	0.329
L	-0.062	-0.09	0.346
M	0.144	-0.05	0.539

Table 4. Improving ratio of detectability

Subject	P <sub>12</sub> (%)	P <sub>23</sub> (%)	P <sub>31</sub> (%)	P <sub>L</sub> (%)
A	31.65	32.01	18.5	32.15
B	23.85	22.23	23.41	23.85
C	15.83	33.74	37.76	37.76
D	18.06	25.67	16.14	25.67
E	16.07	29.01	30.11	30.11
F	15.29	27.88	25	27.88
G	22.17	31.08	26.65	31.08
H	19.45	21.41	29.38	29.38
I	16.14	15.78	24.23	24.23
J	26.59	24.51	25.32	26.59
K	25.61	26.51	19.66	26.51
L	25.98	26.43	19.38	26.43
M	22.7	25.8	15.94	25.8

## 5. CONCLUSION

The improvement method of the detection probability by using multiple modalities condition was proposed in this paper. It is seen from the experiment that the value of the correlation coefficient among modalities is strongly dependent on the subject. The proposed technique will be able to apply the automation for inspection. Because, this technique can give the threshold value of the modality by the detection probability under the condition of multiple modalities, besides it evaluates the detection ability of the panel like the fixed quantity.

## REFERENCES

- Chin, R.T., et al., 1982. Automated visual inspection: A survey, *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol.4, No.6, pp.557-573.
- Drury, C.G., 1972. The effect of speed of working on industrial inspection accuracy, *Applied Ergonomics* Vol.4, pp.2-7.
- Hayashi, H. et al., 2007. An evaluation method of human property using psychometric function in sensory test, *J. Chinese Institute of Industrial Engineering* Vol.24, No.4, pp.259-267.
- Madsen, H.O., Krenk, S. and Lind, N.C., 1986. *Method of Structural Safety*, Prentice Hall.
- Morawski, T.B., Drury, C.G. and Karwan M.H., 1992. The optimum speed of visual inspection using a random search strategy, *IIE Transactions*, Vol.24, No.5, pp.122-133.
- Nakayasu, H., Seya, Y., Nakagawa, M. and Miyoshi, T., 2007. What does a correlation between human responses to different attributes mean?, *Fechner Day 2007(Proc. 23rd Annual Meeting of the International Society for Psychophysics)*, 5pages in CD-ROM.



# SUPPORTING THE PARTICIPATION IN SOCIAL LIFE OF FRAIL OLDER ADULTS IN A ONE-DAY CLINIC

Suzanne Kieffer<sup>1</sup>, Christian Swine<sup>2</sup> and Benoît Macq<sup>1</sup>

*Université catholique de Louvain (UCL)*

<sup>1</sup>*ICTEM – Place du Levant, 2, 1348 Louvain-la-Neuve, Belgium*

<sup>2</sup>*IRSS – Clos Chapelle-aux-champs, 30, bte B1.30.16, 1200 Woluwe-Saint-Lambert, Belgium*

## ABSTRACT

Ageing is associated with an increasing risk to acquire limitations in the daily life activities which is associated with restrictions in the participation in social life. The traditional solutions to address this issue are mainly offered by the health care system, from home care, to hospital admission and in some cases institutionalization in nursing homes. They could be supported and somehow at least in part replaced by Information and Communication Technologies (ICT). *Keep-In-Touch* has been elaborated in this context as a touch-based assistive technology for the older adults. The development method combined user-centered design together with agile development. A six-month pilot study has been undertaken in a one-day clinic in order to efficiently support the development method and evaluate the operational feasibility of the project. First results are promising. The participants judge the system as an effective and simple intergenerational communication tool. They also expressed very positive judgments about the accessibility of the user interface. Consequently, the same approach should be adopted to enrich the system with new functionalities.

## KEYWORDS

Ageing well, assistive technology for older adults, user-centered design, agile development, pilot study, accessibility.

## 1. CONTEXT AND MOTIVATIONS

Ageing is associated with an increasing risk to acquire limitations in the daily life activities (Lang et al, 2007) which is associated with restrictions in the participation in social life (Restrepo and Rozental, 1994). The traditional solutions to address these issues are mainly offered by the health care system, from home care, to hospital admission and in some cases institutionalization in nursing homes. Although these are tested, effective, and operational solutions, they could be supported and somehow at least in part replaced by Information and Communication Technologies (Piniewski et al, 2010). Ambient Intelligence, and especially the smart environment concept, provides a good means to support and facilitate the daily activities of vulnerable persons. Technological innovations, such as ubiquitous computing, sensor networks, wearable devices, tend to contribute to autonomy and well-being, by enhancing safety at home, maintaining social network, mental and physical stimulation (Hine et al, 2005; Cook, 2007; Abascal et al, 2008; Casas et al, 2008).

However, despite the gradual growth of computers (see the following websites: <http://pointerware.com/>, <http://www.elderlycomputer.com/>, or <http://www.eldy.eu>) and interface design guidelines (AgeLight, 2009; Sears and Jacko, 2009a) for the older adults, the digital divide remains significant (Sears and Jacko, 2009a) in particular because of the lack of accessibility of user interfaces. A recent statistical study carried out in Belgium in June 2010 about their use of ICT (sample: 321 persons;  $\alpha=5\%$ ) showed that only 33% of the persons aged 65 and over say that they are interested into ICT *versus* about 50% of the persons aged between 55 and 64 (GeoConsulting, 2010).

Not to mention that in the specific case of frail older adults, the presence of a caregiver is required not only to provide stimulation to the use of the computer, but also to provide assistance during the interaction.

The Keep-In-Touch (KIT) project has been elaborated in this context, as both an assistive technology for the older adults and an interactive tool supporting the caregiver activity, with the following motivations:

Economic and social: decreasing the cost of elder care. Better homecare services and facilitating caregiver's activity should lead to the demand-pull for homecare services together with decreasing the cost of hospitalization (Garrett and Martini, 2007).

Societal: reducing the digital divide. Better accessibility and usability to ICT should lead the older adults to become digital citizens, and consequently surf on the internet, create blogs, use social networking, share knowledge..., and keep in touch with relatives.

Operational: improving the quality of care. Better coordination thanks to customized interactive solutions should lead to the improvement of the integration of care (i.e., to be assured that the patient get the appropriate service, at the right time, in the right place, and from the good person), together with a better continuity of care (i.e., the care are coherent and time-connected).

## 2. KEEP-IN-TOUCH

KIT is an integrated interactive touch-based solution for the older adults providing assistance in daily-life activities, supporting personal autonomy and well-being, and maintaining social cohesion. The solution is easily accessible through an extra-large screen supporting multi-touch interaction. Such a device has been especially chosen because of the following reasons. First, its large resolution may improve the comfort during visual tasks. Second, the touch interaction may reduce the cognitive load for direct manipulation tasks, in particular in comparison with the *keyboard + mouse* paradigm.

A first set of interactive functionalities, *KIT Comfort & Assistance*, has been elaborated especially for the comfort and the assistance of the older adults. The functionalities include: easy access to internet, games, telephone with quick access to emergency numbers, contacts, emails, photos, and files (see Figure 1). More details about the user interface and how it was designed, implemented, and tested will be provided during the oral presentation.



Figure 1. Internet user interface.

A second set of functionalities, *KIT Care*, has been elaborated in order to improve the coordination and the collaboration between the caregivers with and around the older adults. The caregivers are natural caregivers (i.e., family and friends) as well as paid caregivers. The set of functionalities includes a calendar, a communication pad, and easy access to the Resident Assessment Instrument (Morris et al, 1999). The tools are shared between the older adults and their caregivers in compliance with the protection of the private life policy, and have been elaborated in order to improve the communication with and around the older adults, and as a result, to improve the coordination and the integration of care (Sears and Jacko, 2009a).

There are two major innovation aspects, which can be highlighted according to the set of functionalities considered. First, *KIT Comfort & Assistance* provides the end-users with easy access to ICT. By taking into account at different time frames in the life-cycle the needs, expectations, and limitations of the end-users, but also their degree of frailty, the usability of the system has been constantly improved (i.e., overall spatial organization, large components and fonts, easily identifiable components, focus + context visualization, explicit sequences of sub-tasks, multimodal hints). Second, *KIT Care* provides the network of caregivers with the computerized version of a professional paper tool: the communication pad. It is the book where the caregivers share their observations and evaluations about the older adults, describe their activities with her, and attract attention to a potential problem, such as functional decline, loss of appetite, or dementia.

### 3. DEVELOPMENT METHOD

The development method combined user-centered design together with agile software development (Shore and Warden, 2007).

User-centered design employed: knowledge elicitation interviews, paper mock-ups, and focus-group meetings. Domain- and task-relevant knowledge was collected early in the life-cycle thanks to knowledge elicitation interviews. Domain-expert users included professors in geriatric medicine, general practitioners, nurses, health and social professionals. They have been questioned thoroughly about the domain, the task series involved in their activity, their needs and their expectations with the goal to implement the collected information in the system. The equipment used was paper notes and video recording. Paper mock-ups have been implemented according to the recommendations of the experts; in particular regarding the priority of the functionalities to be developed and their content, the interaction scenarios which would best support and fit the activities of the end-users, the size of interactive components and fonts. They have been submitted during focus-group meetings to the panel for improvement and validation.

Agile software development method was adopted in order to develop the prototype, considering the following crucial needs: the collaboration between the experts and the development team, the predictable evolution of the requirements and the solution (see Figure 2), and the rapidity in providing the delivery.



Figure 2. From the first mockup (col a.) to the actual solution (col c.) – keyboard (line1) and user interface (line 2).

### 4. PILOT STUDY

A six-month pilot study focusing on KIT Comfort and Assistance only (see section 2) involving eight frail older adults and one speech therapist has been undertaken in a one-day clinic with the following objectives: first efficiently supporting the development method (i.e., evolution of requirements, bug reporting, usability tests, acceptance evaluation), and second evaluating the feasibility of the project by analyzing how the participants behaved. The pilot study involved eight patients aged 82 (average) with different degrees of frailty (see Table 1 below) rated according to the Edmonton Frail Scale (Rolfston et al, 2006). The computers used were multi-touch all-in-one PC (brand: MSI, model: Wind Top AE2220) equipped with a 21.5" (16:9), 1920 x 1080, full-HD multi-touch screen.

Twice a week throughout the pilot study, the eight participants have been solicited to attend a 2-hour collective session (four participants at a time, one computer each). They were supervised by the caregiver while they performed specific tasks such as web research, email, contact list management, game, etc.

Considering the degree of frailty of participants (see Table 1 below) and their limitations due to ageing (i.e., cognitive loss, cognitive disorientation, memory loss, tremor, etc.), the usability study has been limited to the evaluation of their effectiveness only. Effectiveness refers to the achievement of goals, versus efficiency which refers to the productivity. The effectiveness (i.e., yes or no the end-user can achieve their goals) has been evaluated for the following tasks: log on to the system, web research, add a new contact to the contact list, writing and sending emails, check the weather, set preferences.

Table 1. Degree of frailty of the patients rated according to the Edmonton Frail Scale. Scores: 0-3: no frailty; 4-5: mild frailty; 6-8: moderate frailty; 9-17: severe frailty.

Frailty domain	Patient 1	Patient 2	Patient 3	Patient 4	Patient 5	Patient 6	Patient 7	Patient 8
Cognition	2	1	1	2	1	1	0	2
General health status	1	0	0	2	1	0	1	1
Functional independence	1	1	2	2	1	1	2	2
Social support	0	1	0	0	1	1	0	0
Medication use	1	1	0	1	2	0	1	1
Nutrition	0	0	1	0	0	0	0	0
Mood	1	0	0	1	1	0	1	1
Continence	0	0	1	1	0	0	1	0
Functional performance	2	1	1	2	1	2	2	2
Totals (score /17)	8	5	6	11	8	5	8	9

The first results are promising. Each participant's feedback was collected through a monthly interview conducted by the caregiver right after the session. All of them judge the system as an effective and simple intergenerational communication tool. They also expressed very positive judgments about the large display, the multi-touch interaction, and the contents (internet, photos, and games). According to their caregiver, they perceive KIT as an opportunity to acquire new skills, to be "trendy", and to feel valued. However, moderate to severe frailty (see Table 1, six out of eight patients), sometimes together with dementia (i.e. memory problems and disorientation), makes learning difficult for the patient and wearying for the caregiver: even basic tasks such as finger typing need to be repeated several times until they become familiar. The caregiver plays a key role in spurring and providing assistance to the frail older adults using a computer. Meanwhile, the collaborative use of a computer strengthens the relationship patient-caregiver.

On the other hand, the pilot study allowed following the agile development method. To exemplify this point, let us consider the virtual keyboard. Its usability has been evaluated for months by asking the participants to send emails. It has been simplified and reorganized in order to improve the comfort and the performance of participants while typing text. Complementary to the AZERTY keyboard presented (see Figure 2 line 1), an ABC keyboard has been developed to further facilitating the typing activity. The additional keyboard is used by six out of eight participants.

In a near future, KIT Comfort & Assistance will be enriched with a simplified text editor. The text editor will be used in order to support the edition of the newspaper of the nursing home related to the one-day clinic with the purpose of even more promoting the participation in social life.

## 5. CONCLUSIONS

Keep-In-Touch is a touch-based assistive technology for the older adults that effectively supports the intergenerational communication (i.e., it helps the frail older adults to keep in touch with relatives and caregivers), promotes the participation in social life, and help them ageing well. The development method adopted combined user-centered design together with agile development. It results in the significant involvement of end-users which helps both reducing development cost (time and team) and improving usability (Sears and Jacko, 2009B). Considering the tangible benefits, the same approach should be adopted for the development of KIT Care. However, the drawbacks of agile methods are technical problems, bugs, or too frequent evolutions of the user interface, which may distract the end-users, lead them to a feeling of failure, or even incite them to stop to their participation to the study.

## ACKNOWLEDGEMENT

The research presented in this paper has been financed by the Walloon Region within the First Spin-Off "Keep-In-Touch" project (816788).

## REFERENCES

- Abascal J. et al, 2008. AmbienNet: an intelligent environment to support people with disabilities and elderly people. *In ASSETS 2008*, October 13-15, 2008, pp. 293-294.
- AgeLight, 2009. Interface Design Guidelines for Users of All Ages. In <http://www.agelight.com/>. AgeLight LCC.
- Casas R. et al, 2008. User Modelling in Ambient Intelligence for Elderly and Disabled People. *In Miesenberger K. et al. (Eds.): ICCHP 2008*, LNCS 5105, Springer-Verlag Berlin Heidelberg, 2008, pp. 114–122.
- Cook D.J., 2007. Providing for Older Adults Using Smart Environment Technologies. *In IEEE USA Today's Engineer Online*, 5(7), 2007.
- Garrett N. and Martini EM, 2007. The Boomers are Coming: a Total Cost of Care Model of the Impact of Population Aging on the Cost of Chronic Conditions in the United States. *In Dis Manag*, 10(2):51-60.
- GeoConsulting, 2010. Market study report: the interest of the elderly in ICT in Belgium. GeoConsulting SPRL.
- Hine, N. et al, 2005. Modelling the Behavior of Elderly People as a Means of Monitoring Well Being. *In Lecture Notes in Computer Science*, Volume 3538/2005, 151, DOI: 10.1007/11527886\_32, 2005.
- Lang PO. et al, 2007. Loss of Independence in Katz's ADL Ability in Connection with an Acute Hospitalization: Early Clinical Markers in French Older People. *In European Journal of Epidemiology*, 22(9):621-630.
- Morris JN, et al, 1999. Scaling ADLs within the MDS. *In Journals of Gerontology: Medical Sciences*, 54(11):546-553.
- Piniewski B. et al, 2010. Empowering Healthcare Patients with Smart Technologies. *In Computer*, 43(7): 27-34.
- Restrepo HE. and Rozental M., 1994. The Social Impact of Aging Populations: some Major Issues. *In Social Science & Medicine*, 39(9):1323-1338.
- Rolfston DB. et al, 2006. Validity and Reliability of the Edmonton Frail Scale. *In Age and Ageing*, 35(5):526-529.
- Sears A. and Jacko JA. (Eds), 2009a. *Human-Computer Interaction. Designing for Diverse Users and Domains*. CRC Press: Boca Raton, FL.
- Sears A. and Jacko JA. (Eds), 2009b. *Human-Computer Interaction. Development Process*. CRC Press: Boca Raton, FL.
- Shore J. and Warden S., 2007. *The Art of Agile Development*. O'Reilly Media.
- Takahashi PY. et al, 2010. A Randomized Controlled Trial of Telemonitoring in Older Adults with Multiple Chronic Conditions: the Tele-ERA Study. *In BMC Health Services Research*, 10: 255.

# A MULTIDISCIPLINARY REVIEW: IMPACTS ON DESIGN FOR AGE-RELATED COGNITIVE IMPAIRMENTS

O. K. Burmeister

*Centre for Research in Complex Systems, School of Computing and Mathematics  
Charles Sturt University, Locked Bag 588, Boorooma Street, Wagga Wagga, NSW 2578, Australia*

## ABSTRACT

Research on ageing is being conducted across multiple disciplines. It is important to keep abreast of the wider ageing research, with a view to how findings in other disciplines can impact upon human computer interaction (HCI). Research in psychology on age-related memory losses, in psycho-social areas on biological losses that arise with increasing age, and in biomedicine Alzheimer's related animal experiments, all has potential to impact interface design. HCI researchers need to remain cognisant of advances in other fields. This review examines the possible design implications that arise from such research, and consequently suggests further avenues for HCI research.

## KEYWORDS

Ageing, Alzheimer's disease, human computer interaction, memory, seniors.

## 1. INTRODUCTION

Human computer interaction (HCI) began as a bridging of many disciplines, including computer science, psychology and others, and has emerged as a discipline in its own right. Amongst the many HCI studies, there have been numerous investigations of ageing (Pfeil and Zaphiris, 2010, Zaphiris et al., 2005, Sayago and Blat, 2010), including usability and accessibility guidelines (Jaeger and Xie, 2008), mobile devices (Kurniawan, 2008, Ziefle and Bay, 2004), pervasive computing (Shankar, 2008), various interfaces that aid senior users of information and communications technology (ICT) (Hawthorn, 2007, Rice and Alm, 2008), and a few that have focused on issues of cognition (Sayago and Blat, 2010, Burmeister, 2010, Czaja et al., 2006, Hiltz and Czaja, 2006). In HCI the value of multidisciplinary research is well recognised. For instance, one key area of HCI is that of participatory design, which is in effect a set of practices that draws on multiple disciplines, including graphic design, software engineering, psychology, sociology and political science (Muller, 2003). Similarly, in the area of ageing, it is important to recognise the value of multidisciplinary approaches. There have been at least four discoveries in other disciplines that could benefit interface design for seniors, particularly in relation to age-related cognitive impairments. This paper reports on those four studies, and on how they impinge on HCI. Thus future HCI research directions are revealed, that offer new possibilities for HCI designers, to meet the needs of senior users of ICT.

Amongst the multidisciplinary discussion are two discoveries in psychology, one that separates cognition into two main types, the other is a related finding revealing that, certain types of age-related memory impairments can be compensated for, with the provision of contextual markers. A third discovery that has the potential to impact HCI research is psycho-social literature on the lifespan theory 'selective optimization with compensation' (SOC) and the concept of assisting seniors through technology, to compensate for age-related functional losses. Finally, a study reported this year in biomedicine revealed that mice with Alzheimer's disease (AD), have difficulty to find 'home', when other mice without that disease, had no such difficulties. Given that it appears that pharmacological solutions to AD remain distant, this suggests that certain types of navigational aids could assist seniors, who suffer from the form of dementia known as Alzheimer's disease.

The paper reviews the findings in each of these aforementioned four studies. Each such discussion ends with an implications section, to demonstrate how the findings could be applied in a HCI setting. The paper ends with overall conclusions regarding the importance of multidisciplinary research, to the advancement of HCI research.

## **2. COMPENSATING FOR FUNCTIONAL LOSSES**

Selective optimization with compensation (SOC) is a psycho-social lifespan theory of ageing that has gained prominence in the last two decades. SOC researchers (Baltes, 1997, Marsiske et al., 1995, Baltes and Smith, 2002) found that, from an evolutionary perspective, selection pressures operated in the first half of an individual's lifespan, to ensure reproductive fitness, as well as effective parenting behaviours. SOC envisages the progression to successful ageing as going beyond reliance on evolutionary biology, to cultural influences. Successful ageing therefore, is less a function of biology, and more a function of increasing the culturally-based resources available to people to help them find supportive compensations for biological losses. It is an integrative approach that incorporates multiple factors affecting seniors as they age. It combines individual, social and institutional perspectives. SOC recognises that people cannot do everything themselves and that they compensate for this by relying on external supports in those areas (Baltes, 1997, Marsiske et al., 1995).

In other words, as Vanderheiden (Vanderheiden, 1997, 1994) foreshadowed over a decade ago, HCI can have an important role in helping seniors to overcome the functional losses of ageing. More recently, ethnographic HCI research has revealed that when the social capital available to seniors is increased, their enjoyment of ICT-based social interaction and their overall well-being increases (Pfeil and Zaphiris, 2010, Xie, 2008). However, such findings appear as independent studies, which do not look beyond HCI, such as to psycho-social studies of ageing to explain why this might be the case. By combining HCI and other ageing research, it may be possible to extend technological aids to seniors beyond current assistive technologies and accessibility guidelines. Such things are important and provide important compensatory steps toward better resourcing seniors, but more needs to be done.

## **3. DECLINES PREDOMINANTLY AFFECT MECHANICAL COGNITION**

SOC studies concerning the cognitive abilities of seniors have revealed that change occurs in adaptive capacity in two related areas. Those two areas are referred to as mechanics and pragmatics. Mechanics describes the biological intellectual functions concerning information processing, such as "reasoning, spatial orientation, or perceptual speed", whilst pragmatics describes "cultural systems of inheritance", for example "verbal knowledge (e.g. semantic memory) and certain facets of numerical ability" (Baltes et al., 1999). Furthermore, in relation to mechanical cognition, research with seniors in developed countries has demonstrated that people maintain mental achievement levels until about age 70 (Baltes and Smith, 2002). Then, the older people become, the more seniors experience impairments when they encounter new learning. Baltes and Smith stated that this is particularly evident in the very old, and that sizeable losses in people's ability to learn can occur for people over 85. Thus one aspect of mechanics, concerns the ability of seniors to process new information.

How might this impact on HCI design? One implication is that care needs to be taken with people over the age of 70. The older they are, beyond that age, the more difficulties they will have in adapting to software upgrades, changes to hardware platforms and new interfaces. The reason for that is that, all such changes require users to relearn to use their technology, but for such seniors, the new learning involved can be particularly difficult. For example, in a study on mobile phone use by seniors, Kurniawan exposed seniors to a complex new interface, and concluded that seniors cannot cope well with complexity (Kurniawan, 2008). However, had that HCI researcher been cognisant of the psycho-social cognition research above, an alternate conclusion might have been that it was not complexity, but new learning, that seniors found problematic in the use of such a new mobile phone interface.

## **4. CONTEXTUAL INFORMATION COMPENSATES FOR DECLINES IN EPISODIC MEMORY**

The term 'semantic memory' referred to in the quotation by Baltes, Staudiner and Lindenberger, above, is employed by cognitive psychologists to refer to a person's store of factual knowledge. Craik (1999) claimed that seniors experience little or no decline in semantic memory performance, but "substantial age-related

declines in episodic memory performance ... (which) ... may be taken as a further proof of the existence of at least two separate memory systems.” He further stated that episodic memory can be supported with the provision of contextual information.

A few HCI researchers have already begun to explore the impacts of two memory systems on HCI design. For instance, the memory types referred to here as semantic and episodic, have also been referred to as crystallized and fluid intelligence (Czaja et al., 2006), respectively, and were recently employed in a study of web use by seniors (Hanson, 2009). Hanson described crystallized intelligence as the ability to use skills, knowledge, and experience, whereas fluid intelligence is the capacity to think logically and solve problems, independent of acquired knowledge.

Unlike mechanical performance (episodic memory / fluid intelligence), pragmatic abilities (semantic memory / crystallized intelligence) are not affected by increasing age, although there is some evidence of decline in the very old (Baltes et al., 1999, Craik, 1999). It appears to this author that until now most HCI ageing research involving cognition, has treated all cognition as a single entity, rather than taking into account cognitive psychological research, that for over a decade has shown that instead, there are at least two distinct types of cognition. Understanding the difference, and designing interfaces that compensate for mechanical losses, could advance HCI further, in this important area.

## **5. NAVIGATIONAL AIDS TO FIND HOME COMPENSATE FOR ALZHEIMER’S RELATED LOSSES**

A recent biomedicine study further confirms the importance of navigational aids that provide seniors with contextual information. Hamlin, Windels, Sah and Coulon (2011) investigated previous work which had discovered that the loss of basal forebrain cholinergic neurons (BFCN) was a feature of Alzheimer’s disease. They “selectively ablated BFCN with saporin conjugated to a p75 antibody in c57B16 mice” and found that, like non-lesioned mice, lesioned mice did not lose gross motor skills. However, unlike mice not so lesioned, analysis of associative memory “revealed that lesioned mice showed comparative levels of freezing behaviours in both contextual, and classical conditioning paradigms” (Hamlin et al., 2011). Furthermore, “when lesioned mice were subjected to a passive place avoidance paradigm where they had to learn to avoid a spatial region (60° zone of an 800mm diameter open field, devoid of cues) that had been previously paired with a mild footshock (0.4mV) they performed poorly, unable to recall the spatial location of the shock zone” (Hamlin et al., 2011). They conclude that these trials with mice show that there was no decline in learning and memory, but instead, there was a decline in navigational ability, when there was an absence of spatial cues. In particular, the lesioned mice had difficulty going back to previous locations and notably, finding ‘home’.

One implication of such research, assuming that it translates to humans, is that HCI designers need to provide seniors not only with contextual information, particularly where new learnings are involved, but that they also need to provide all seniors, even when no new learning is involved, with easy ways to navigate backwards and particularly to ‘home’.

## **6. CONCLUSION**

At its heart HCI is not a unitary discipline, but is a multidisciplinary field. This review has drawn on four non-HCI studies, to show that future HCI research that involves age-related issues of cognition, can benefit from a multidisciplinary understanding of cognition and seniors.

The research reviewed here supports the view that increasing the social capital available to seniors, through technology, can compensate them for functional declines, particularly mechanical cognitive declines. Furthermore, in relation to mechanic cognition, new learning is difficult for some people over the age of 70 and, for increasing numbers, the older they become. Finally, in relation to memory deficiencies that can increase with ageing, the provision of contextual information and particularly the ability to navigate backwards and to ‘home’, is an important design consideration, and one that warrants further research within the HCI community.



## ACKNOWLEDGMENTS

I am indebted to David Sloan and Sergio Sayago, both from the School of Computing, University of Dundee, Scotland, for their review and comments on the first draft of this paper.

## REFERENCES

- Baltes, P. B. (1997) On the incomplete architecture of human ontogeny: Selection, optimization, and compensation as foundation of developmental theory. In Baltes, P. B. & Brim, O. G. J. (Eds.) *Life-span development and behavior*. New York, Academic Press.
- Baltes, P. B. & Smith, J. (2002) New frontiers in the future of ageing: From successful ageing of the young old to the dilemmas of the fourth age. *Valencia Forum* April 1-4 ed. Valencia, Spain.
- Baltes, P. B., Staudinger, U. M. & Lindenberger, U. (1999) Lifespan psychology: Theory and application to intellectual functioning. *Annual Reviews Psychology*, 50, 471-507.
- Burmeister, O. K. (2010) Websites for Seniors: Cognitive Accessibility. *International Journal of Emerging Technologies and Society*, 8, 99-113.
- Craik, F. I. M. (1999) Age-related changes in human memory. In Park, D. & Schwarz, N. (Eds.) *Cognitive aging: A primer*. Philadelphia, Psychology Press.
- Czaja, S. J., Charness, N., Fisk, A. D., Hertzog, C., Nair, S. N., Rogers, W. A. & Sharit, J. (2006) Factors predicting the use of technology: Findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE). *Psychology and aging*, 21, 333.
- Hamlin, A., Windels, F., Sah, P. & Coulson, E. (2011) The role of basal forebrain cholinergic neurons in spatial memory. Barcelona.
- Hanson, V. L. (2009) Age and web access: The next generation. *W4A2009 - Technical*. Madrid, Spain.
- Hawthorn, D. (2007) Interface design and engagement with older people. *Behaviour and Information Technology*, 26, 333-341.
- Hiltz, S. R. & Czaja, S. J. (2006) Introduction to the special issue on information systems for an aging society. *ACM Transactions on Computer-Human Interaction*, 13, 309-312.
- Jaeger, P. T. & Xie, B. (2008) Developing online community accessibility guidelines for persons with disabilities and older adults. *Journal of Disability Policy Studies*, 20, 55-63.
- Kurniawan, S. (2008) Older people and mobile phones: A multi-method investigation. *International Journal of Human-Computer Studies*, 66, 889-901.
- Marsiske, M., Lang, F. R. & Baltes, P. B. (1995) Selective optimization with compensation: Life-span perspectives on successful human development. In Dixon, R. A. & Baeckman, L. (Eds.) *Compensation for psychological defects and declines: Managing losses and promoting gains*. Hillsdale, NJ, Erlbaum.
- Muller, M. J. (2003) Participatory design: The third space in HCI. In Jacko, J. A. & Sears, A. (Eds.) *The human-computer interaction handbook*. Mahwah, New Jersey, Lawrence Erlbaum Associates.
- Pfeil, U. & Zaphiris, P. (2010) Investigating social network patterns within an empathic online community for older people. *Computers in Human Behavior*, 25, 1139-1155.
- Rice, M. & Alm, N. (2008) Designing new interfaces for digital interactive television usable by older adults. *ACM Computers in Entertainment*, 6, 1-20.
- Sayago, S. & Blat, J. (2010) Telling the story of older people e-mailing: An ethnographical study *International Journal of Human-Computer Studies*, 68, 105-120.
- Shankar, K. (2008) Pervasive computing and an aging populace. *Journal of Information, Communication and Ethics in Society*, 8, 236-248.
- Vanderheiden, G. C. (1994) Application software design guidelines: Increasing the accessibility of application software to people with disabilities and older users.
- Vanderheiden, G. C. (1997) Design for people with functional limitations due to disability, aging, or circumstances. In Salvendy, G. (Ed.) *Handbook of Human Factors and Ergonomics*. New York, John Wiley & Sons.
- Xie, B. (2008) Multimodal computer-mediated communication and social support among older Chinese internet users. *Journal of Computer-Mediated Communication*, 13, 728-750.
- Zaphiris, P., Kurniawan, S. & Ellis, R. D. (2005) Web and Aging: Challenges and Opportunities. *Universal Access in the Information Society*, 4, 1-2.
- Ziefle, M. & Bay, S. (2004) Mental models of a cellular phone menu. Comparing older and younger novice users. *MobileHCI 2004*. Springer-Verlag.

# A CONTROL AND TRAINING SYSTEM FOR REHABILITATION DEVICES

Patrick Patterson and Eric Sinzinger  
*Texas Tech University  
 Lubbock, Texa, USA*

## ABSTRACT

A major shift in how electronic devices can be controlled is proposed. The proposed project, if successful, will allow the control of any electronic device (i.e. prosthetic limb, computer) by use of neural and minor muscular inputs. We are developing an adaptive algorithm capable of choosing a best set of signals/signal locations, on an individual basis, from the large number of possible combinations available as initial control solutions. As the user becomes more adept with controlling the device, this personalized adaptive algorithm will learn which control signals provide the most effective control. Similar to learning to walk, coordination of the various sensors will adapt, improving overall control.

## KEYWORDS

Device control, Rehabilitation, Control algorithm

## 1. INTRODUCTION

The need for devices and systems that can complement or replace human functions is an ongoing challenge. Loss of a limb due to injury or disease reduces an individual's ability to interact with the environment. Injury to the spinal cord or brain may cause permanent loss of nearly all voluntary motor function and sensation. Even those with full use of their faculties may need skills and actions supplemented in different device-control situations.

There are three primary approaches to improving human-device interaction:

1. *Use of muscle movement to control/activate a device.* In this approach a body part physically moves to control the device, such as devices activated by the tongue (Salem and Zhai, 1997), the head (Origin Instruments, Grand Prairie, TX), or the eye's gaze (Foulds, Arthur, and Khan, 1996), and even speech recognition. Consistent control of a monitorable body part is required for use of these systems.

2. *Use of biosignals to monitor and control systems.* These electrical signals may be read from the human body in a number of places such as individual muscles or muscle groups (electromyographic or EMG) or the eye (electro-oculographic or EOG). Both the EMG and EOG use electrical side effects of muscular control and usage. Using such systems, people with full muscle control can move their own limbs to control an interface, while those without full muscle control must retain some nerve function in the limb, amputated extremity, or in a remote part of the body (Guger, et al., 1999). Confounding the accurate monitoring of these neural outputs is the continuous presence of interference from sweat, emotions, fatigue, and the need for continuous concentration during activities. These interfering factors have thus far made device control based on these signals difficult unless performed under a very restrictive set of conditions.

3. *Use of brain-computer interfaces (BCI).* BCI interfaces are often referred to as think-to-activate interfaces. Capturing an analogue electrical signal directly from the brain without an additional biosignal carrier has the advantage of being a suitable interface for persons with no motor control. Medical technology provides new methods of displaying biological states (such as real-time blood pressure and body temperature). With more refined feedback, it is clear a person *can* control even biological functions previously considered involuntary (Lubar and Lubar, 1984). This process, biofeedback, has been used to gain conscious control over many aspects of the autonomic nervous system. The BCI approach to device control uses EEG waves as the signal source. Modeling the EEG presents difficulties, most of which can be attributed to complications with the signal or with determining model configurations. This interface seems

applicable to a wide range of users as it incorporates the more easily produced and controllable EMG signals, but still requires the user to concentrate on producing specific brain activity to monitor control frequencies by concentrating on a varying bar graph. Even if the control capabilities are excellent, this need for intensive concentration and visual distraction hampers the overall control process in real-world activities. Researchers, such as Pfurtscheller (2010) are developing new hybrid systems, but currently they function primarily for off/on, left/right, binary selection. Even given these shortcomings, the characteristics of such a hybrid system concept will serve as the basis for our investigations.

A large variety of classification methods have been used to process biosignals. For example, artificial neural networks, such as the adaptive logic network (Kostov and Polak, 2000) and the standard multi-layer perceptron, which can be supervised or unsupervised, have been used. RBF networks have the advantage of requiring a relatively small training set as well as not getting stuck in local minima. Alternative approaches are linear discriminant analysis (LDA) classifiers or linear vector quantization (LVQ) classifiers. LDA have the advantage over LVQ of requiring less training trials to configure the classifier for real-time use (Pfurtscheller, et al, 2000), and does not require any preconfigured parameters. However, these methods do not adapt well to substantial changes in the operating environment, as would be required for real-time device control. In some cases, if the classification certainty drops the system loses its ability to accurately respond- a potentially disastrous situation for the user of a mobile robot, prosthesis, or wheel chair.

## 2. OUR APPROACH

There is substantial knowledge about the characteristics of motor activity and biosignals. Approaches 2 and 3 above for human-device interaction can be categorized as human assisted neural device techniques. Each of these approaches has had limited success in a variety of applications, with BCI think-to-activate systems being the most difficult to design/use. Few methods can successfully use these techniques in a designed system to control an electronic device under a wide variety of circumstances.

The objective of this work is to investigate how a combination of biosignals and muscle movements can be used to control an electronic device, with an eventual goal of remote sensing. The proposed hybrid system will be flexible enough to allow a user to gain control over any properly equipped device through the interpretation of these input signals. The system will continually improve its performance, requiring minimal or no preliminary training to function properly. The system will also be designed for use in a variety of environments. A brief summary of the approach is presented below.

### 2.1 Input Signal Acquisition and Correlation

Our initial approach is to use a basic system with characteristics as discussed in the following. As innovations and improvements occur, changes will be made to this first generation system with the intent to design a less cumbersome on-person/in-environment signal collector.

We will use a variety of biosignals, similar to the approach of Grychtol (2010), to determine the optimum signal combination for a particular user. With this collection of input data (a user's biosignals), we will generate a population of functions, each of which take some or all of the collected data as input. All functions will be correlated with the response to the subject's attempted action. Those functions with the lowest correlation will be deleted and we will then generate variations of the functions with high correlation, using a shotgun-and-focus approach called the "Great Deluge" algorithm. When a numerical combination of inputs is successful for one subject, it will be added to a database of known successful combinations. This database will support the initial learning algorithm and should reduce the need for training for the system to function for each new subject. We can also mine this database to evaluate types and placements of sensors for later design studies for more diverse applications such as harsh geographic terrain and extreme weather conditions.

Support vector machines (SVMs) are to be investigated as an alternative approach to mapping biosignals to the desired functionality (Hsu and Lin, 2002). SVMs combine machine learning with statistical pattern theory to provide binary classification on a given set of inputs. For each desired control action, a unique SVM is to be constructed that determines for a given input whether to perform the given action. Given a set of inputs with training data, the SVM embeds the inputs in a higher dimensional space that allows a

hyperplane to divide the true and false actions. Having multiple SVMs has the benefit of not only providing a stable classification for notoriously noisy signals, but also allowing multiple desired actions to be detected and performed simultaneously.

## 2.2 System Optimization

This system will constantly monitor and interpret input signals, process them to determine meaning, and respond accordingly. It is essential that the system continually improve its algorithms as use continues to ensure real-time, accurate feedback with the chosen electronic device. This will be accomplished using two optimization algorithms running simultaneously: one for mapping input signals into output responses and the other ensuring the first system is robust, so fast and effective interaction is always maintained.

Processing the signals to output responses for a particular subject beyond the initial learning algorithm will be accomplished with machine learning tools such as evolutionary computing, fuzzy computing, and neuro-computing. These techniques are generally considered to be bottom-up tools, that is, order and structure emerge from an unstructured beginning. Included among the CI tools we will use are neural networks, fuzzy logic, genetic algorithms, self-organizing maps, and learning vector quantization.

A second methodology will be developed to design the structure of the input processing system. It is envisioned that this will involve the use of evolutionary algorithms that are excellent at finding global optimal solutions to a problem while also providing a range of solutions. We will take advantage of both of these characteristics. First, achieving the overall best solution for the arrangement of the input processing system, will allow a network to adapt to a particular subject in real-time. As a subject uses the system, continual network modification will be taking place improving the system's response and performance. In this manner, we effectively will be designing an algorithm to optimize another algorithm. A current "best" solution for the input network will always be maintained. As a subject generates inputs, the network will respond to them. When the continual optimization process determines a better system configuration, it will replace the current best configuration and used for all subsequent input responses.

This continual algorithm optimization will create a highly adaptable learning algorithm able to respond in real time to various physical and emotional states of the participant (e.g., fatigue, anger, and happiness) that could otherwise hamper correct action identification.

## 3. SUMMARY AND FUTURE WORK

Results of our initial work is being used to create a prototype device/system that, in real-time, will have the capability of adapting to users and their states by interpreting muscular and neural signals to control a motorized/computer device such as a myoelectric prosthesis, vehicle, robot, or computer interface. In brief, subjects will do a series of activities with sensors collecting performance data. The data will be used to improve and optimize the control algorithm for that individual. Like biofeedback, this mapping of the biosignal pattern to the action will be accomplished subconsciously. If a mistake is made, we will track the corrective measures such that even mistakes will improve the performance of the device-human interaction. A prototype of the system will then be tested with a wide variety of users to determine robustness. The system will auto configure and support machine learning in real-time for each user. This can be used to address, for example, the population of mobility-impaired individuals, a category that is increasing due to an aging population, better medical care, and increasing numbers of military amputees.

## REFERENCES

- Salem C., Zhai Z., 1997. "An isometric tongue pointing device," In *Proceedings of the CHI '97 Conference on Human Factors in Computing Systems*, pages 538 - 539, New York.
- Foulds R., Arthur J., and Khan A., 1997. "Human Factors Studies in Eye Movements Related to ACC Head Movement Studies," *Rehab R&D Progress reports* 1996, vol.34, 155-156.
- Grychtol, B., et al. "Human behavior integration improves classification rates in real-time BCI," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 4, pp. 362-368.

- Guger, C., et al, 1999. "Design of an EEG-based braincomputer interface (BCI) from standard components running in real-time under Windows." *Biomed. Technik*, vol. 44, pp. 12-16.
- Lubar, J.F., & Lubar, J.O., 1984. "Electroencephalographic biofeedback of SMR & beta for treatment of attention deficit disorders in a clinical setting," *Biofeedback & Self-Regulation*. 9, 1-23.
- Kostov A. and Polak M., 2000. "Parallel man-machine training in development of EEG-based cursor control," *IEEE Trans Rehabil Eng* 8: 203–205.
- Pfurtscheller G., et al, 2010. "Self-paced operation of an SSVEP-based orthosis with and without an imagery-based "brain switch:" a feasibility study towards a hybrid BCI," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 4, pp. 409-414.
- Pfurtscheller G., et al, 2000. "Brain oscillations control hand orthosis in a tetraplegic," *Neurosci. Lett.*, vol. 292, no. 3, pp. 211–214.
- C. Hsu and C. Lin, 2002. "A comparison of methods for multiclass support vector machines," *IEEE Transactions of Neural Networks*, Vol. 13 (2), pp. 415-425.

# TR-AID ADVANCED SEARCH: TAXONOMY DESIGN FOR DATA VISUALIZATION AND INFORMATION RETRIEVAL INTERFACE

Massimiliano Gusmini<sup>1</sup> and Anjula Garg<sup>2</sup>

<sup>1</sup>*Via Ippocastani 8, 20152 Milano*

<sup>2</sup>*European Commission, Joint Research Centre  
Institute for the Protection and Security of the Citizen  
Via E. Fermi, 2749 I-21027 Ispra (VA) Italy*

## ABSTRACT

This paper discusses the design process of a tailored multi-index, multi-source information retrieval system (TR-AID). It provides an efficient comprehensive visualization of information for donors and controllers dealing with development and humanitarian aid and associated data. The GUI presented in this paper, especially the Advanced Search panel, makes it possible to browse large amounts of information without losing its contextual relevance. Index integration has been achieved by a cognitive approach. A detailed analysis (also called mapping technique) produced a taxonomy which is used as a set of filters, to provide high quality data retrieval to amplify cognition for the users in an advanced relational context.

## KEYWORDS

Cognitive interfaces, multi-user, multi index, multi-source IRS

## 1. INTRODUCTION

The European Union and the Member States need to coordinate their policies on development cooperation and consult each other on their aid programmes in order to promote the complementarity and efficiency of their action. TR-AID (Transparent Aid) supports this mandate for European coordination by making aid related information available from a single platform. Specifically, in the context of the agenda on aid effectiveness, TR-AID concentrates on enhancing the coordination between EU donors by sharing data. This increases complementarity and coherence of donors' action by making such data available in a consistent and understandable form even for those who may not be experts in the field of aid.

The work described in this paper was carried out at the Joint Research Centre (JRC), of the European Commission (2010). The project TR-AID is an initiative driven by the European Commission with the primary objective of facilitating the sharing of aid funding data amongst donors, so that the data can be used in a decision making capacity across all levels of the funding lifecycle.

The TR-AID project includes a web based system developed to support sharing of information amongst donors with the aim of using aid funds most effectively and facilitating the detection and prevention of misuse of aid funds. TR-AID combines data from multiple sources into its own database, and allows the publication of comprehensive, timely and detailed information about humanitarian and development aid, in a form that is easy to access. This is done by allowing the users to search the database for information relating to projects and organisations (linked with recipients, implementing partners and donors), both via a textual and also a graphical interface. The system has been set up in such a way that with a minimum of search criteria input, a wealth of information can be obtained from the system. Currently the database holds information about 40,000 organisations, 1.2 million projects and related data. It also stores other important information such as sectors in which projects are being carried out.

## 2. BODY OF PAPER

### 2.1 Multi-user, Multi-Index and Multi-Source Information Retrieval System

The context of handling large amounts of data required the investigation of the potential of information retrieval based on multiple indexes, sources and user requirements. In this context, EU donors incorporating different types of users are integrated through an adaptive interface so as to enable browsing in different types of documents and information, reflecting the complex work environment of a shared information system, with a diverse set of users. The main requirement was to provide high quality data results.

In this paper we describe how we produced the application GUI, especially the Advanced Search panel, which is the focal panel of our information retrieval system. The process of working with heterogeneous parameters of Search Criteria started with analyzing the requirements of different users who have different types of responsibility in managing aid funds. User types include: 1. policy makers (those who make decisions on planning and funding of projects), 2. implementers (those who actually facilitate funding of projects) and 3. controllers and anti-fraud investigators (who check that projects are implemented as originally planned). The tool must handle the complexity of multiple requirements coming from different types of user that have very different objectives and methods according to their activity type.

Multi-index retrieval systems combine heterogeneous types of content and very different logic criteria behind them, such as specific parameters and values. The task was to define categories of items and to place them in a taxonomic scheme. Then we were able to normalize the depth levels of definition, to fit with each type of category of parameters.

Multi-source systems are related to aggregation, normalization and the visual display of data from different databases, data sources, and data fields. We wanted to give the user the feeling of total integration of data. Sources are multi-selectable and switchable. By default all sources are selected.

To be able to control this complexity, we designed a flexible taxonomic structure, able to contain the particularity of several parameters of different human activities, workflows and ways of thinking.

### 2.2 Cognitive Approach for Interface Design: Mapping Matter, Modeling GUI

We started the conceptual work with a cognitive approach in mind. We believe that Information Visualization can amplify the user's cognition (Card, S. K. - Mackinlay, J. D. - Shneiderman, B. 1999). Cognition, meaning as "the process of thought" and as "the acquisition or use of knowledge", has been our main paradigm in the first approach that was intended to allow all our users to be familiar with the contents, subject matter, parameters and their structured category and retrieval paths, in order to make the most effective workflow possible. We think that this was necessary to facilitate comprehension and memorization for the user. In order to give the user this value of usability of a tool, we chose to build a clear tree map taxonomy for selecting search parameters according to visualized criteria. We wanted to model the interface design on the *Mental Image* (Thorndike P.W 1982) of the taxonomy. This is the reason why we avoided hidden classification of contents such as a set of drop-down menus or an input-text based interface.

To do this we had to deal with topics like the users' short-term memory, mental location of items, and the general mapping of the matter. For aspects related to short-term memory we looked at the experiments and the conclusions of Miller (1956) and Baddeley (1986). Miller suggests that seven (plus or minus two) is a good number of items to memorize in the short term. Baddeley later suggests to group items (called chunk) to overcome this limit. During the design process we paid special attention to respect these limits. A first issue in fact, has been the analysis and the identification of the main categories of parameters. We wanted the smallest number possible of category-items at the first level. This was in order to simplify the working memory (Normann 1999) and reduce the impact of complexity as revealing complexity gradually helps the user to approach the knowledge (Maeda 2006).

To enhance cognition of matter and efficiency of the workflow, it is also important to offer the users a Cognitive Map that helps them to memorize: location of items, paths to reach them, and also, workflow processes. For these aspects we looked at research by Thorndike and Heyes-Roth in 1982 on spatial representation. According to this theory, humans use two main types of spatial knowledge for interacting with space: "route" representation, that refers to specific paths; "survey" representation, that is a global map overview, a perspective from above. It has been shown (Pazzaglia, Cornoldi e De Beni 2000) recently that

certain particularly relevant reference points, called “landmarks” (not necessary in a spatial context) such as “route” or “survey”, can be useful in the exploration of an environment. Landmark representation consists of identifying relevant places and using them to make spatial relationships. Paths, maps and landmarks, are ways of offering the user the Mental Map (Nicoletti R.- Rumiati R 2006), the most comprehensive way of displaying the matter, especially in very complex contexts. Some interesting reflections during the design process were about the importance of a natural relationship between the mental map and the actual location of contents, when there is a need to memorize information. We found interesting suggestions from ancient cartography and how cartographers used visual metaphors to amplify memory, knowledge and cognition, in physical and mental environments (Mangani G. 2006).

### 2.3 The Taxonomy of the Search Parameters define Interface and Workflow

The visual organization of data by structural relationships enhances the recognition of patterns and increases the effectiveness of a task (Card, S. K. - Mackinlay, J. D. - Shneiderman, B. 1999). Categories are the core of taxonomy design. They respond to the principle of “cognitive economy” because they allow us to maximize information and distinguish them from each other. ( Nicoletti R.- Rumiati R 2006). Rosch and Mervis (1975) demonstrate how categories, in this case displayed propositionally, are structured in two main dimensions: *vertical* (main category) and *horizontal* (sub-category), and how this display can help the economization of the comprehension effort. Other useful criteria for the taxonomy design is what Bersalou (1983) called *category “ad hoc”*. The interesting theoretical aspect of this type of category is about the inclusion of items that do not have any characteristics in common but they can be structured gradually as other taxonomic categories. This category has been frequent in TR-AID contents with specific and sophisticated parameters.

The *structure principle* (Larry Constantine 1999) has been another fundamental paradigm in our design process because it is strictly related to the activity of contents classification in an organized taxonomy. It says “*Design should organize the user interface purposefully, in meaningful and useful ways based on clear, consistent models that are apparent and recognizable to users, putting related things together and separating unrelated things, differentiating dissimilar things and making similar things resemble one another.*” We structured the taxonomy as a tree-map, we then empirically located the categories, by doing several tests of logic processing in order to obtain coherent positions of items. The result of this process gave the most efficient ways of browsing and selecting our parameters, without redundant or missing options.

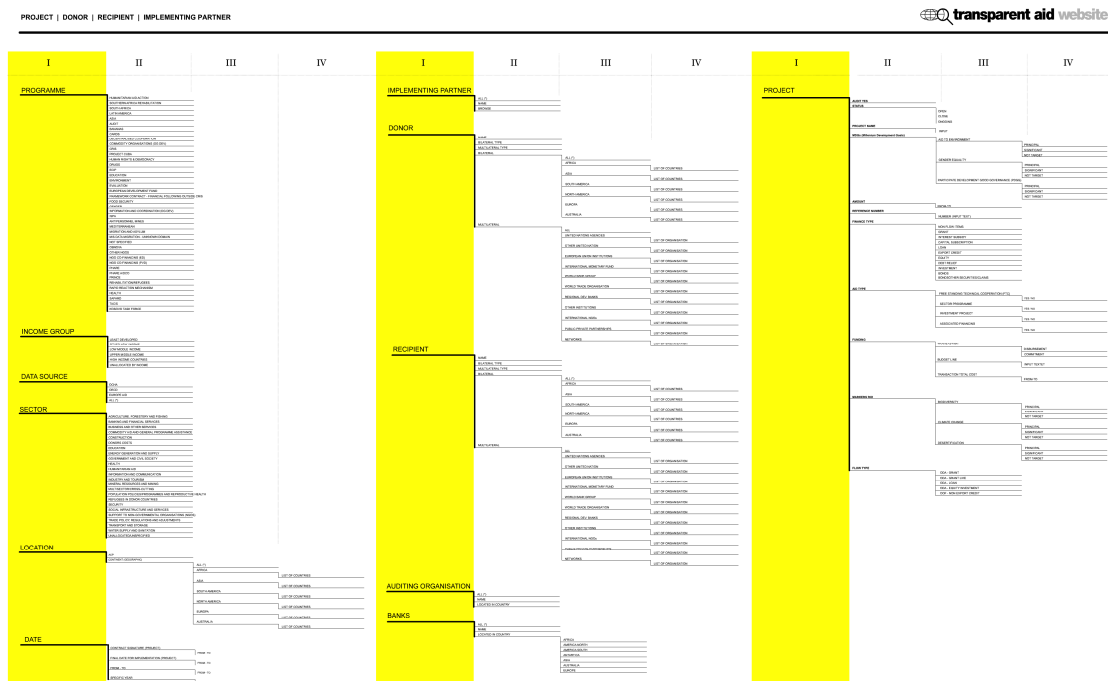


Figure 1. TR-AID taxonomy. Four level depth. Yellow columns (Gray in black and white) are first level



## 2.4 Interface and Workflow Description

The TR-AID website interface fully respects the Principles of User Interface Design (Larry Constantine-Lucy Lockwood 1999) as does the TR-AID Advanced Search panel. TR-AID website is a “Focus plus Context Interface” (Furnas G.V. 1986), all in one window application, conceptually organized in three main areas: Search (basic and advanced), Retrieve and Basket (Fig. 2).

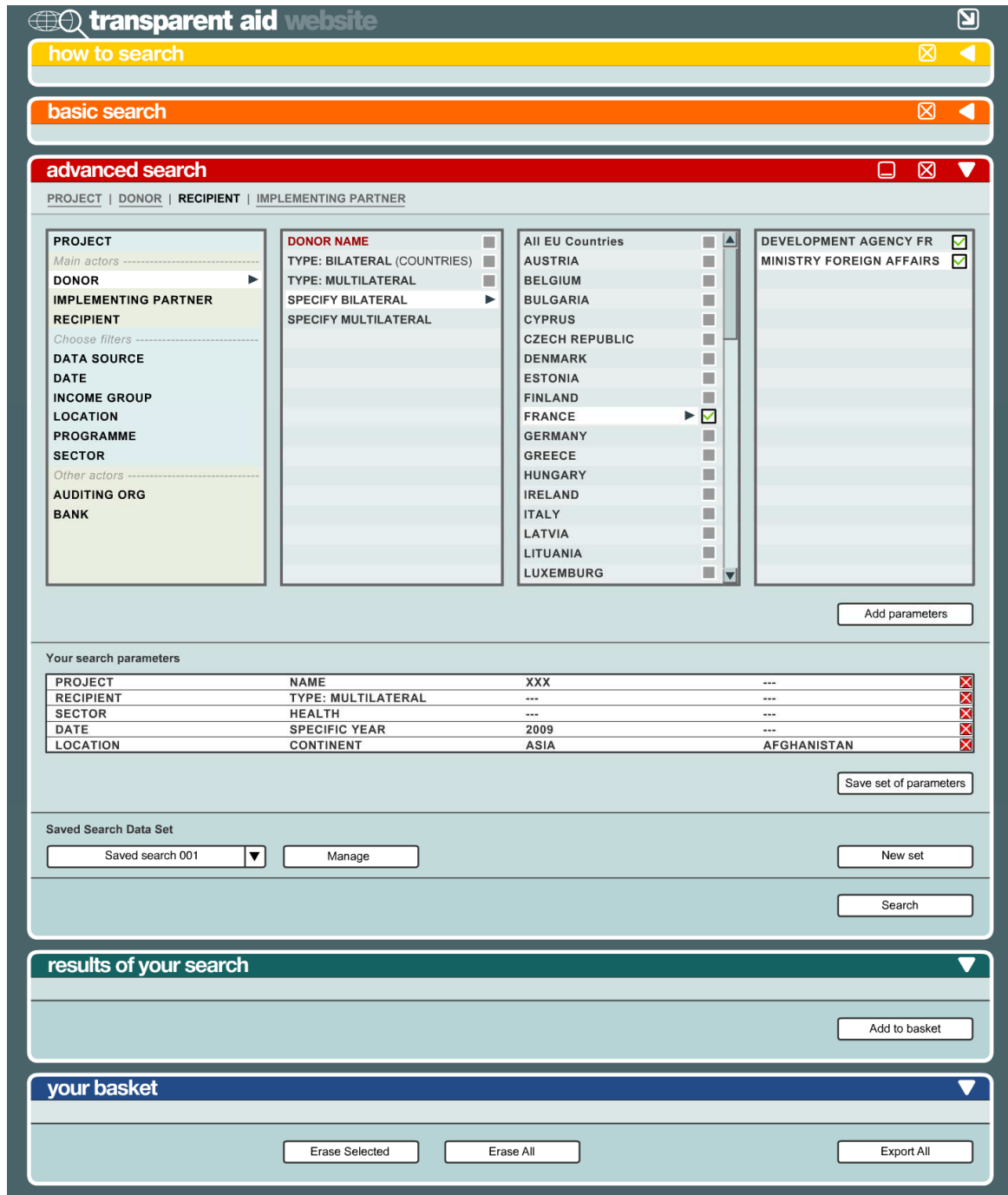


Figure 2. TR-AID GUI. Example of transposition of the taxonomy into the Advanced Search GUI panel. The table below is the summary of selected parameters. Search parameters for filtering data are multi selectable in order to restrict results.

The Interface, modeled on this conceptual partition, displays these areas as collapsible panels. In the interface we added a Guideline panel above the other panels, and it can be closed or collapsed after a few sessions of work. The Graphic Interface lets the user customize these areas according to the type of work.

Here we briefly describe all panels from top to bottom. At the top, the Guideline panel displays generic information on the work process and helps understand the workflow. The Basic Search panel which is a text input search module, is a useful tool when the user knows specific names or keywords. Next is the Advanced Search panel, as described below, the most efficient way to find high quality data for projects, organizations, donors, recipients and implementing partners.

Next is the Retrieve panel (called *Results of your search*), which is a table panel with expandable rows that allow users to see result items and details. After a search session it lists all the results with 25 items on each page. Once the user has checked the results, some (or all) can be selected and placed in the Basket. The Basket panel is the final panel, it allows the user to export search results and use them in an external analytical tool. This last step closes the TR-AD Search workflow with saving and exporting data that has been retrieved.

A user can search the database for projects funded by different donors without specifying any parameters. However, in this case, the user is presented with 1.2 million projects which need a huge effort to study and understand. Alternatively, a more informed user can use the taxonomy as a set of Search parameters, and obtain focused and more manageable results. After the Search parameters have been selected, we can activate the search engine by clicking the Search button.

The major issue in the design process has been the transposition of the taxonomy displayed graphically (**Fig. 1**), into the GUI of the Advanced Search panel. We transposed the taxonomy into an interactive panel with conventional GUI elements for software applications. This is a set up panel for Search parameters, with multi-selectable options that will act as filters. There are two functional areas in this panel: the selection of search parameters and the summary of selected parameters. The search parameters selection area is for browsing and selecting search filters based on the structure of the taxonomy, translated into a Column View interface (also called Miller columns) that fit well with the tree map. The area below is the Selected Parameter Summary, a list that summarizes the selected parameters for search filtering in the current session. This set of parameters can be saved and used for future work sessions. Although there can be a maximum of four levels of depth, most of the contents classification remains within two levels. The column view is very intuitive and helps the user to visualise categories, parameters and ways of data-filtering in order to amplify cognition. Browsing a tree map taxonomy through columns, makes it easier to memorize locations of parameters. Another task was the hybridization of two types of common GUI modules: Miller columns and Input-text field. This was also designed to match our users' interests in investigation of specific organization or project names.

As an example of workflow we report three different ways of querying data from TR-AID according to our type of users and their objectives of searching data.

User case 1 (Donor): List of closed projects, for Education in Angola between 2006 and 2009 financed by Luxemburg in OECD data source. The search parameters set up starts by identifying and selecting the subject, in this case PROJECT. Then we define the SECTOR (Education), DATE (From 2006 – To 2009), LOCATION (Geographic, Angola), PROJECT (status closed), DONOR (Luxemburg), DATA SOURCE (OECD).

User case 2 (Antifraud investigator): List of all Multilateral Recipient in Afghanistan in 2009 related to the project "alpha", on Health activities, financed by France government. The subject in this case is the RECIPIENT. Then RECIPIENT (MULTILATERAL), PROJECT name ("alpha"), SECTOR (Health), DATE (2009), LOCATION (Geographic, Afghanistan), DONOR (France).

User case 3 (Antifraud investigator): List of implementing partners organisation related with the project name "beta", in the year 2010, financed by all EU Donors in all Data sources. This case would follow this path. Subject: IMPLEMENTING PARTNER, then specify the PROJECT name ("beta"), then the year (2010) and then the donor (all EU States).

### 3. CONCLUSIONS

In our experience the approach and method implemented here, dramatically improved the effectiveness of data retrieval. Our challenges were: to enhance cognition of the underlying information, to facilitate memorization of work processes, to offer an intuitive navigation, and to reduce the number of operations for the user to obtain high quality results. Using the taxonomy as a conceptual interface helped us to design a multi-requirement, multi-index and multi-source Information Retrieval System. Both users and implementers can benefit from the display of category-branches, which imposes a logical flow on the process of option selection, and so highlights missing options and encourages assessment and tuning of parameters. This logical flow also improved the optimization of the system parameters, as illogical and incoherent options were easily identifiable. This helps the overall evaluation of system efficiency.

### ACKNOWLEDGEMENT

We would like to thank: Valdemar Blazevic, Tony O'Grady, Andrius Sunelaitis, Francisco Diaz and Christina Goodwin for their valuable feedback.

### REFERENCES

#### Books:

- Card, S. K., Mackinlay, J. D., Shneiderman, B., 1999. *Information Visualization. Using vision to think*. Morgan Kaufmann Publisher, USA pp. 6-28
- Constantine L. 1999. *Software for Use: A Practical Guide to the Essential Models and Methods of Usage-Centered Design*. MA: Addison-Wesley USA.
- Coren, S., Lawrence M., Ward, James T. Enns (1999). *Sensation and Perception*. Harcourt Brace. p. 9.
- Löwgren, J. and Stolterman, E., 2004. *Thoughtful Interaction Design*. Massachusetts Institute of Technology, USA
- Kosslyn S.M. 1994. *Image and brain: The resolution of the imagery debate*. Cambridge, MA, MIT Press USA
- Mangani G. 2006. *Cartografia morale. Geografia, persuasione, identità*. Franco Cosimo Panini Publisher. ITA
- Maeda, J. 2006. *The Laws of Simplicity*. Massachusetts Institute of Technology, USA

#### Links:

Constantine L. - Lucy Lockwood. *Principles of User Interface Design*, source Wikipedia  
[http://en.wikipedia.org/wiki/Principles\\_of\\_User\\_Interface\\_Design](http://en.wikipedia.org/wiki/Principles_of_User_Interface_Design)

#### Journals:

- Barsalou, L.W., 1983. *Ad hoc categories*, in «Memory and Cognition», 11, pp 211-227.
- Furnas, G.W., 1986. *Generalized Fisheye Views*. In Human Factors in Computing Systems CHI '86 Conference Proceedings, pp. 16-23.
- Miller, George (1956) *The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information*. . In «The Psychological Review», vol. 63, pp. 81-97.
- Rosch, E., 1975. *Cognitive Representations of semantic categories*. In «Journal of Experimental Psychology: General», 104, pp. 192-233.
- Rosch, E. and Mervis, C.B., 1975. *Family resemblances: Studies in internal structure of categories*. In «Cognitive Psychology», 7, pp. 573-605.
- Thorndike, P.W., Hayes-Roth, B., 1982. *Differences in spatial knowledge acquired from maps and navigation in* «Cognitive Psychology, 14, pp. 560-589.

# A COMBINED AGILE METHODOLOGY FOR THE EVALUATION OF WEB ACCESSIBILITY

Firas Masri and Sergio Lujan Mora

*Department of Software and Computing Systems, University of Alicante, Spain*

## ABSTRACT

To assure and certify the fulfillment of web accessibility guidelines (*WCAG 1.0 & 2.0*) and to guarantee accessibility to all disabled users, various accessibility evaluation methods have been proposed, basically classified in two types: qualitative methods (*analytical and empirical*) and quantitative methods (*metric-based methods*). As no method by itself is enough to guarantee full accessibility, many studies advice to combine these methods between each other in order to guarantee better results. Some recent studies present combined evaluation methods between qualitative methods only, thus leaving behind the great power of metrics that guarantee objective results and task diversity. To achieve this goal, the current paper proposes a systematic combined agile accessibility evaluation method based both on qualitative and quantitative evaluation methods. This proposal presents an evaluation methodology that combines between the essential analytical evaluation methods and the empirical user test methods. Finally, WAB metric (*Web Accessibility Barrier*) is included to summarize objectively the final results and amplify the use of this method to cover all types of evaluations tasks, like validating, certifying and comparing processes.

## KEYWORDS

Web accessibility, Empirical evaluation methodology, User test.

## 1. INTRODUCTION

Nowadays the web is present in all fields of life, from access to information and service web pages to management of legal documents by means of electronic public administration (*e-government*).

This makes that users will be a heterogeneous public, with different abilities and disabilities (visual, hearing, cognitive & motorize impairments). These characteristics represent a huge challenge if we hope to provide universal access to all possible users, specially if the intention is to fulfill web accessibility guidelines WCAG 1.0<sup>1</sup> and WCAG 2.0<sup>2</sup> (*Web Content Accessibility Guidelines*) of the WAI (*Web Accessibility Initiative*) or the other national and international guidelines and laws (Requisiti tecnici, 2005; Section 508, 2007; UE Commission, 2002) which regulate and protect the right of disabled users to access information.

To assure web accessibility, several studies have suggested numerous evaluation methods (Brajnik, 2006; Bühler, 2006; Vigo, 2007) as a means to verify, measure and certify the fulfillment of the accessibility guidelines and therefore to supply full accessibility to disabled people. Currently, there are two types of evaluation methods: the qualitative methods (analytical and empirical) and the quantitative methods.

The qualitative methods have been the most used until now, specially the analytical ones, which are characterized by their low cost and ease of use. Automatic evaluation tools such as ATRC,<sup>3</sup> WAVE 4.0<sup>4</sup>, EvalAccess 2.0<sup>5</sup>, TAW<sup>6</sup> and Cynthia Says<sup>7</sup> have been the pioneers and the most well-known, due to its usability, ease of use and its quick results, although they are not the final and complete solution. The other

<sup>1</sup> Available at: <http://www.w3.org/TR/WCAG10/>.

<sup>2</sup> Available at: <http://www.w3.org/TR/WCAG20/>.

<sup>3</sup> Web Accessibility Checker. Available at: <http://achecker.ca/checker/index.php>.

<sup>4</sup> WAVE 4.0 - Web accessibility evaluation tool. Available at: <http://wave.webaim.org/>.

<sup>5</sup> EvalAccess 2.0 - Web service tool for evaluating web accessibility. Available at: <http://supt07.si.edu.es/evalaccess2/index.html>.

<sup>6</sup> TAW - Web accessibility test. Available at: <http://www.tawdis.net/>.

<sup>7</sup> Cynthia Says evaluation tool. available at: <http://www.contentquality.com>

analytical evaluation methods, which are based on the manual heuristic inspection of code, do not guarantee full accessibility (Brajnik, 2008); it depends largely on evaluators' experience and the adopted guidelines. On the other hand empirical methods are generally more expensive, but more accurate, because they clearly show the most catastrophic accessibility faults. User test is the most reliable and complete one (Masri & Lujan, 2010a).

The quantitative methods help to understand, control and improve the final product (Fenton & Pfleeger, 1998), thus its main goal is to assure the quality results and monitor the accessibility level by establishing values and summarizing results. These methods and due to their nature aren't sufficient enough to assess accessibility and evaluators can't depend only on them.

As consequence of the above-mentioned situation related to evaluation methods and because no method on its own would guarantee the detection of all accessibility barriers, some studies like (Lopez, 2010; Villegas, 2010), considering W3C and researchers recommendations, starts to apply in their evaluations process a combined evaluation method based completely on qualitative methods (analytical and empirical methods). These works ignored totally the inclusion of the important quantitative methods to their methodology.

Due to this situation, and addressing the growing need of providing combined evaluation methods, arises the idea to present both kind of methods in a clear documented evaluating methodology, with a clear user – centered orientation, specially that there is no homogeneous and unified practice in this field (Masri & Lujan, 2010b), and those that already exist are not complete enough as to guarantee the coverage of all kinds of evaluations tasks from on side and does not provide a clear and objective results from other side.

The methodology presented in this paper combines the essential analytical methods with the empirical user test method and concludes with the WAB metric (*Web Accessibility Barrier*) (Parmanto and Zeng, 2005) that summarize objectively the results; thus permitting the coverage of all evaluation tasks.

## 2. ACCESSIBILITY EVALUATION METHODS

Qualitative methods include two essential types of evaluation methods: the analytical and the empirical ones. The analytical methods are: standard expert's revisions (conformity revisions), the automatic tools and barrier walkthrough method. On the other hand, the empirical methods are: user tests, subjective revisions, screen techniques (Masri & Lujan, 2010a).

As for the qualitative evaluation methods finality, they are used to undergo formative evaluations (identification of the list of problems) during the development phase; and summative evaluations (validation and comparison) in the final phase of the product and after the final users have used it. These methods estimate the accessibility of an interface so as to validate it. The results will always show the descriptions of the failure modes, defects or even solutions and recommendations for the developers.

On the other hand, quantitative methods are: Failure Rate metric, WEBQEM (*Web Quality Evaluation Metric*), WAB metric (*Web Accessibility Barrier*), UWEM (Unified Web Evaluation Methodology), A3 (Aggregation Metric), WAQM (*Web Accessibility Quality Metric*) and T1 metric. These methods helps evaluators to monitor and improve accessibility levels, also permits to summarize results objectively, their results can be used to compare quality among web pages, or to track quality improvement in the quality assurance process (Vigo et al, 2007; Freire, 2008; Sirithumgul, 2009).

### 2.1. Comparison between Web Accessibility Evaluation Methods

Evaluation methods whether qualitative (analytical & empirical) or quantitative have advantages and disadvantages when evaluating web accessibility.

Analytical methods are characterized by their great capacity of identifying a wide range of diverse problems for diverse audiences, apart from their ability of marking the exact violations of the adopted guidelines. However, they are criticized for requiring skilful evaluators and for not distinguishing between the important from unimportant web accessibility problems (Brajnik, 2008). Analytical methods have also proved methodologically weakness when qualifying the gravity of identified problems (Petrie & Kheir, 2007). A comparison between automatic tools also shows quite contradictory results (Thatcher et al, 2006; Diaz & Cachero, 2009).

Empirical methods are more exact when qualifying a web site; they discover the more catastrophic accessibility mistakes in a real time, especially if they are applied in the correct context taking into account the specific characteristics of the web accessibility explained in (Masri and Lujan, 2010a). Many studies (Dey, 2004; DRC, 2004) have shown that empirical methods and specially user tests are the unique ones that detect and qualify the severity of the real mistakes faced by users. They have shown also that 45% of accessibility difficulties emerge with the application of user tests to disabled people after being evaluated with software and other methods. However, they are criticized for their elevate cost due to necessity for equipped labs and real users.

On the other hand, the quantitative metrics are being used in large-scale evaluations processes; only the metric for WEBQEM did not. WAB, UWEM, A3, WAQM and T1 metrics have shown a great correlation between their results (Freire, 2008; Sirithumgul, 2009), for that reason, it is not possible to state which metric could be more effective in general cases. Each metric may be more suitable for different projects, according to their needs. In this sense, in order to help the definition of good metrics, (Daskalantonakis, 1992) it is needed to identify important characteristics of useful software metrics. According to Daskalantonakis, software metrics must be: (1) simple to understand and precisely defined; (2) objective; (3) cost effective; and (4) informative (ensure that changes to metric values have meaningful interpretations).

From all the above mentioned characteristics, we consider that a combined accessibility evaluation method must contain: (1) user test method (due to its reliability and effectiveness); (2) quantitative metric (to control and monitor accessibility results and cover all tasks types). Particularly, in our proposal, we consider that WAB metric presented by (Parmanto and Zeng, 2005) is the best metric that fits our approach due to its simplicity, objectivity and task coverage, beside, it has a fixed defined barrier weight that correspond to WCAG levels (A, AA & AAA). Also, it is considered inclusive because it includes all user groups.

### 3. THE COMBINED METHODOLOGY FOR EVALUATING WEB ACCESSIBILITY

When we talk about accessibility evaluation, first of all, we must think directly what our principal objective is and what we know about it. When assessing accessibility, we can face three different situations; the first is: evaluating a site during development in order to see if it fits all accessibility guidelines; the second is to validate or compare its compliance with accessibility guidelines after being adapted by developers and posterior to its launch; however, in the third situation, we don't know anything about the site we want to evaluate and we are obliged to start from the initial steps to check the real state of the site, which means from the analytical evaluation methods.

The first situation can be resolve by applying the common analytical automatic evaluation methods, while in the second case, it's more complicated and the best practice is to apply an empirical evaluation test, while in the third situation, many evaluators lose time to determine how to assess it. So taking into consideration the three possible cases, we propose two steps for the three goals commonly or separately.

- The analytical evaluation step
- The empirical evaluation step

#### 3.1 The Analytical Evaluation Step

This step consists of two stages:

- **The pre analysis stage:** illustrate a rapid and clear idea about the site that is going to verify.
- **The automatic evaluation stage:** Helps the evaluator to have a complete idea about the site understudy and helps to determine whether the site is qualified to continue to the next step directly or not.

##### 3.1.1 Pre Analysis Stage

This stage is composed of the following main activities:

- Localize and verify the identity of the site under evaluation to make sure that's the correct one
- Compare the actual version of the site with oldest version using [www.archive.org](http://www.archive.org), thus avoiding unexpected mistakes and having an idea about the sites evolution.

- Navigate the site in order to have a more complete idea about the used design.
- Apply software tools like Web Developer Tool to revise HTML cleanliness and detect image volume used in the design.
- Apply text and sound browsers like WebbIE and others to verify if information is accessible and equivalent to that provided by a graphical user interface browser.

### 3.1.2 Automatic Evaluation Stage

This stage is composed of the following main activities:

- Validation of mark-up language using *HTML Validation Service of W3C*.
- Validation of CSS style sheet using *CSS validation service*.
- Automated Accessibility evaluation using TAW or other available automatic evaluator software.

## 3.2 The Empirical Evaluation Step

Current legislations like (UE Commission, 2002; Section 508, 2007) establish priority/level AA as mandatory for all public and electronically products; so it was decided to consider two levels of test:

1. Test of level 1.
2. Test of level 2.

Test level 1 has to do with the conformity degree A and AA, where the biggest users groups will access information without troubles. On the other hand, test level 2 is related to the conformity degree AAA that is a higher level of fulfillment in which all users groups will access without any kind of barriers. It is suggested doing the test in just one day in order to reduce costs and increase the test control and systematization.

### 3.2.1 Users and Staff

The following users and staff are involved in an empirical evaluation:

- Users with different disabilities (Vision, hearing, motor & cognitive impairments), 6 are enough according to (Nielsen, 2000), 1 or 2 per each disability. Users must have some experience with internet and assistive tools.
- Facilitator, one person is enough to explain users the sequence of steps.
- Assistant or observer. An observer per disability is enough to guarantee the understanding and observation of the process.

### 3.2.2 Room and Equipment

It can be carried out in a small office or laboratory supplied with two computers for each selected disability. The computer must have the specific devices and software for the disability in question, for example: voice browsers, screen readers, a screen magnifier program, Jaws program, Synthesizer + Tiflowin software, etc. For more information about devices and software see (Pernice & Nielsen, 2001).

### 3.2.3 Test Planning

**Test plan:** Accessibility guidelines are prepared, classified and revised according to the test level (A & AA for test level 1 and AAA for test level 2). The web site must be studied to localize functionalities and the supposed potential barriers. For better results, it is advisable to assign tasks focusing on the fulfillment of the most essential functionalities with an implicit orientation towards each disability.

**Material Preparation:** The preparation of all tasks to be developed is undergone based on the test planning. The obtained material must be attractive and not exhausting for the user, with questions that put the site accessibility at stake. For example:

1. Browsing through the main menus, localizing the most attractive links (suggesting some of them and covering most of them).
2. Following a link predetermined by the evaluator (image-link).
3. In case there is a multimedia, asking for a summary of the multimedia subject.
4. Locating some essentials functionality and ordering user to access through the keyboard.
5. Filling up a form or sending a message through the page.

6. In case the site has sale service or donations options, asking the user to select one product and proceed to the purchase.

7. In case the site allows the formal step of legal documents, request the user to proceed to fill in the application form and document's formalities.

### 3.2.4 Required Tests

**Pilot Test:** It is performed to make sure that everything will turn out just as it has been planned and that the media work properly. The user does not have to be necessarily a disabled person.

**Final test:** The test is performed by the selected users in a familiar environment, free of tension. Users are asked to use the protocol, think aloud.

### 3.2.5 Qualification Methodology

A proposal of this property should be done through an approach of observation on one hand and a metric that establishes a value and summarizes the results on the other hand. The analysis system must be as it follows:

- **User's interview:** The interview results are compared with the corresponding observer's notes.
- **Recording analysis:** The recording is analyzed in order to find errors omitted by the observers.
- **Data analysis:** Errors found are analyzed and classified, taking into account the priorities and its corresponding checklists.
- **Conclusion:** A final report on the data analysis results is drawn up
- **Evaluations Metric:** The evaluation metric criteria come to be the test barometer (measure) that allows providing a concrete result. The formula results as it follows (Parmanto and Zeng, 2005):

$$WAB = \frac{\sum_{j=1}^T \sum_{i=1}^n \left( \frac{b_{ij}}{B_{ij}} \right) (W_i)}{T} \quad \text{Where,}$$

$b_{ij}$  Represents the real errors and barriers in a priority

$B_{ij}$  Represents the potential barriers of each page in such priority

$W_i$  Represents the gravity of the real error inverse to each priority (1-2-3)

$T$  Represents the number of pages of the whole site

This metric will be applied to the final report of the concluded data analysis results, it is considered best result those that is close or equal to 0.

## 4. CONCLUSION

The proposed combined method is clearly exposed step by step as a clear guide for web accessibility evaluations, the methodology ends with a clear and objective qualification method based in WAB score weighted metric formula that resume and validate the obtained results. This methodology characterized by its systematic and objective results with a user centered evaluation focus; it is systematic because all of its steps are developed formally and under observation, and considered objective because it resumes the results by mean of an evaluation metric and not subjectively with human opinion. The next step of our research will be the application of this method to evaluate different Spanish public and commercial websites, thus improving improper derived mistakes and including more user experience to this combined evaluation solution.

## REFERENCES

- Brajnik, G., 2008. Beyond conformance: the role of accessibility evaluation Methods. *In 2nd International Workshop on Web Usability and Accessibility IWWUA08*, Auckland, New Zealand.
- Brajnik, G., 2006. Web accessibility testing with barriers walkthrough. Visited: Feb 2011. Available at: [www.dimi.uniud.it/giorgio/projects/bw](http://www.dimi.uniud.it/giorgio/projects/bw).



- Bühler, C., Heck, H., Perlick, O., Nietzio, A., and Ulltveit-Moe, N., 2006. Interpreting Results from Large Scale Automatic Evaluation of Web Accessibility. In *Computers Helping People with Special Needs*, Springer, and Berlin/Heidelberg.
- Daskalantonakis M. K., 1992. A Practical View of Software Measurement and Implementation Experiences within Motorola. *IEEE Transactions on Software Engineering*, 18(11):998 - 1010.
- Dey, A., 2004. Accessibility evaluation practices—survey, results. Visited: Feb 2011. Available at: <http://deyalexander.com/publications/accessibility-evaluation-practices.html>.
- Díaz, B., Cachero, C. 2009. Fiabilidad de herramientas automáticas para el aseguramiento de la accesibilidad en Web. In *IADIS Ibero-Americana WWW/Internet 2009*, Madrid, Spain.
- Disability Rights Commission (DRC). 2004. *Formal investigation report: web accessibility*. ISBN 0 11 703287 5.
- Fenton N. E., Pfleeger S.L. Software Metrics: A Rigorous and Practical Approach. PWS Publishing Co., Boston, MA, USA, 1998.
- Freire, A. P., Fortes, R., Turine, M., Paiva, D., 2008. An Evaluation of Web Accessibility Metrics based on their Attributes. *SIGDOC'08, September 22-24, 2008*, Lisbon, Portugal.
- López, L. AWA: *Marco metodológico específico en el dominio de la accesibilidad web para el desarrollo de aplicaciones web*. Tesis Doctoral, Universidad Carlos III, Leganés, Madrid. 2010.
- Masri, F., Lujan, S., 2010a. Análisis de los métodos de Evaluación de la Accesibilidad web. *Proceedings of the: Universidad 2010 conference*. Havana, Cuba.
- Masri, F., Lujan, S., 2010b. Test de usuario: Un Método Empírico Imprescindible para la Evaluación de la Accesibilidad Web. *Proceedings of the: 5ª conferencia ibérica de sistemas y tecnologías de información 2010*. Santiago de Compostela, España.
- Nielsen, J., 2000. *Why You Only Need to Test with 5 Users*. Visited March 2011, available at: <http://www.useit.com/alertbox/20000319.html>.
- Parmanto, B., & Zeng, X. M. 2005. Metric for web accessibility evaluation. *Journal of the American Society for Information Science and Technology*, 56(13), 1394-1404.
- Pernice, K., Nielsen, J., 2001. Beyond Alt Text: Making the web easy to use for users with disabilities. Visited: Feb 2011, Available at: <http://www.nngroup.com/reports/accessibility/testing/>.
- Petrie, H., Kheir, O., 2007: The relationship between accessibility and usability of websites. In: *Proc. CHI 2007*, pp. 397–406. ACM, CA
- Requisiti tecnici e i diversi livelli per l'accessibilit`a agli strumenti informatici, 2005. (G.U.n.1838/8/2005), Visited Feb 2011, [www.pubbliaccesso.it/normative/DM080705.htm](http://www.pubbliaccesso.it/normative/DM080705.htm).
- Romen, D., Svanaes, D., 2008. Evaluating Web Site Accessibility: Validating the WAI Guidelines through Usability Testing with Disabled Users, *proceedings: NordiCHI 2008, October 20 to 22, 2008*.
- SEC.508, 2007. ELECTRONIC AND INFORMATION TECHNOLOGY, Visited Feb 2011, Available at: [www.section508.gov/index.cfm?FuseAction=Content&ID=14](http://www.section508.gov/index.cfm?FuseAction=Content&ID=14).
- Sirithumgul, P., Suchato, A., Punyabukkana, P. 2009. Quantitative Evaluation for Web Accessibility with Respect to Disabled Groups.
- (SSWAE), 2011. Spanish Senate web accessibility evaluation. Available at: <http://accesibilidadweb.dlsi.ua.es/?menu=ej-analisis-senado-parte-1> y <http://accesibilidadweb.dlsi.ua.es/?menu=ej-analisis-senado-parte-2>.
- Thatcher, J., Burks, M., Heilmann, C., Henry, S., Kirkpatrick, A., Lauke, P., Lawson, B., Regan, B., Rutter, R., Urban, M., Waddell, C., 2006: Web Accessibility: Web Standards and Regulatory Compliance; Friends of ED.
- UE Commission., 2002. Visited: Feb 2011. Available at: [http://ec.europa.eu/ipg/standards/accessibility/index\\_en.htm](http://ec.europa.eu/ipg/standards/accessibility/index_en.htm).
- Vigo, M., Arrue, M., Brajnik, G., Lomuscio, R., and Abascal, J. 2007. Quantitative metrics for measuring web accessibility. In *W4A '07: Proceedings of the 2007 international cross-disciplinary conference on web accessibility (W4A)*. Banff, Canada.
- Villegas, E., Pifarré, M., Fonseca, D., 2010. Diseño metodológico de experiencia de usuario aplicado al campo de la accesibilidad. *Actas de la 5ª conferencia ibérica de sistemas y tecnologías de información 2010*. Santiago de Compostela, España.

# DEVELOPING E-LEARNING SYSTEMS FOR DEAF EDUCATION: APPLICATION OF EMBODIEMENT THEORIES

Maria Tagarelli De Monte<sup>1</sup>, Daniele Capuano<sup>2</sup>, Maria Roccaforte<sup>3</sup> and Elena Tomasuolo<sup>4</sup>

<sup>1</sup>*Istituto Statale per Sordi di Roma , Via Nomentana, 56, Roma*

<sup>2</sup>*Università Sapienza di Roma, Pictorial Computing Lab Department, Via Salaria 113, Roma*

<sup>3</sup>*Università Sapienza di Roma, Dipartimento di Scienze Documentarie Linguistico-Filologiche e Geografiche  
P.le Aldo Moro, 5 Roma*

<sup>4</sup>*Istituto di Scienza e Tecnologie della Cognizione, CNR, Via Nomentana, 56, Roma*

## ABSTRACT

When addressing linguistic problems of prelingual deaf people, research in the field of deaf education has proven how technology can enhance the quality of content transmission and of student's learning experience. E-learning facilities allows reaching isolated groups of deaf learners and address their pedagogical needs through an immersive and challenging learning environment. We will describe the working framework of the FIRB-VISEL project, aimed at creating an interactive e-learning environment where deaf people can improve their linguistic skills in the local language. In order to address both signers and non signers, we focus on how to convey information and learning strategies without losing track of their learning advancements and their motivation. We found that embodiment theory and its application to education could be beneficial. In this paper, we outline the path we have followed in the development of learning objects based on embodiment theory and their application to the field of deaf education.

## KEYWORDS

Deaf education, embodied cognition, recreational e-learning environments, storytelling, visual interaction, literacy.

## 1. INTRODUCTION

People with pre-lingual deafness encounter major problems in acquiring their national language, even in its written form. In a digitalized world, where a lot of information appears in a textual form, this represents a serious problem.

Within the FIRB-VISEL project we aim to develop a Deaf-centered E-Learning Environment (DELE) for promoting literacy skills and improving access to web-based education in young and adult deaf learners. Our target is composed by prelingual deaf people (either using sign language as their main means of communication or not) who do not possess sufficient linguistic skills to access written information. One of our goals is to design an e-learning environment as visual as possible, therefore respecting deaf people's preferred intact sensorial channel, limiting the use of written text to exercises. In order to achieve this goal, we are relying on the set of approaches coming from the application of embodiment and cognitive linguistics theories, applied to a design work done in team with deaf and hearing researchers. In this way, the user interface and navigation structure generated from this work should truly reflect the needs of the deaf.

In this paper, we will show the results of the first two years of the VISEL project, sketching the first modules of DELE. The discussion is organized as follows: Section 2 presents the literacy problems encountered by deaf people, while our theoretical approach is explained in Section 3; Section 4 describes the DELE structure and Section 5 concludes the paper.

## **2. DEAFNESS AND LITERACY**

All over the world, deaf children and, later, adults experience dramatic difficulties in achieving appropriate receptive and expressive skills not only in oral language but also in written language. The fact that literacy problems appear in all linguistic communities could be seen as a deafness-specific more than a language-related issue. On the other hand, these problems are specific to each culture and each language, and they are not always comparable. In Italian, for example, deaf people show particular weaknesses in the use of free morphology, pronouns, prepositions, articles and the verbal system. It must be emphasized that these problems related to written language do not depend on the speech rehabilitative therapy undergone, but are common to the whole deaf population, both signing deaf and orally educated deaf (Caselli et al., 2006; Fabbretti, Tomasuolo, 2006).

On the other hand, research on the visual skills of deaf people (Muir, Richardson, 2005; Jensen, et al., 2000; Muir et al., 2003) has shown how developed their visual skills are, and how they can be used to convey multiple meanings.

The aim of the VISEL project is to meet the educational needs of the deaf, both deaf learners who prefer to communicate in Italian Sign Language – LIS (LIS-L1) and deaf learners who prefer to communicate in Italian (Italian-L1) - through a strong investment in their visual skills. The e-learning tools that we intend to design and produce is aimed to effectively promote appropriate receptive and expressive written language skills in deaf Italians who study in high schools and/or Universities or who are professionals involved in the education of deaf children and/or in LIS courses designed for hearing, hard of hearing or deaf people.

## **3. EMBODIMENT AND LINGUISTIC STRUCTURES: FEELING AS A MEAN FOR TEXT ACCESSIBILITY**

In developing DELE, we need to consider that, even though our aim is to design a poor-of-text environment, texts *are* the main objects of our didactic model, since the learners attending our virtual courses have to reach a good understanding of texts, in all of their forms. Thus, how can we enhance text accessibility for deaf people?

In our culture, texts are seen as mainly intellectual, mind-concerned materials. As stated in Steidele (2004), the history of reading seems to be characterized by a progressive “loss of the body”, opposed to ancient approaches to reading where people were asked to abandon themselves to the richness of feelings re-created by texts. Such feelings are related to the complex set of bodily actions and responses which emerges when reading is performed: the vibrations of mouth and lips, the motion of tongue, etc. In this way, the symbolic meaning of words was enriched and fully recreated in the reader. In the development of DELE we wish to re-create this experience in our learners building an interactive environment in which physical responses to reading, as well as the general “feeling” conveyed by written language, will be visualized.

The Embodied Cognition paradigm (Langer, 1967; Johnson, 2007; ; Damasio, 2003) supports this view. This theory proposes a holistic and ecological model where humans’ understanding is built starting from the interactions in the environment. As the contact between the subject and the environment occurs, a continuous flow of unconscious emotional responses is generated and, when such responses are strong enough, a qualitative conscious feeling is produced. Repetitive patterns of experience generate basic non-propositional “gestalts” of knowledge, called image schemas.

The entire human cognition starts from such “felt qualities” (Johnson, 2007) - pre-conceptual backgrounds which permeate and build the meaning of our experiences - and is orchestrated through image schemas. Rational concepts and symbolization emerge from these felt qualities as discrete units of meaning extracted from the continuous flow of experience. For example, if we see a landscape we first have a general qualitative impression of the entire environment, and only in a following stage are we able to recognize single elements. The Embodied Cognition theory claims that the first qualitative feeling is a fundamental part in our meaning-making process as much as our high-level, rational and analytic cognitive tools.

This is also true in the case of language, where the application of the embodied cognition paradigm is addressed by cognitive semantics. In this field, the notion of image schema is the most important theoretical notion, constituting the form of representation common to perception, memory and semantic meaning (Gardenfors, 2007).

In the case of the use of verbs, Talmy (1988) convincingly demonstrates that a great deal of our understanding depends on the forces that are involved in the action expressed by the verbs. To “understand” we have to see the complete sentence and to associate the meaning with the context. For example it is quite easy to explain verbs that express movement, like *venire* (“to come”), *andare* (“to go”), but it is nearly impossible to explain their “felt quality” (“affordances”) to foreign learners, meaning the particular use of these kind of verbs in the Italian language. This is also the problem for deaf learners.

In the development of DELE we are seeking to transmit language “affordances” using metaphors of their use. For example, in the case of connective words such as *nonché* (“as well as”), *nonostante* (“although”), etc. we will be using the metaphor of a bridge, which is something that meets two ends, but it’s also something that allows access. In DELE, an animation explains the right position of the connective in the sentence. Symbols, arrows, image schemas, diagrams help in the explanation and can be used for navigation as well as in the activities and exercises, encouraging manipulation of the written language by deaf learners. Learners who are LIS-L1 could be helped by videos in Sign Language in which the concept (the connective in this case) is not translated but explained through visual expression, gesture, examples.

#### 4. VISUAL NAVIGATION, ACCESSIBLE NAVIGATION

The DELE structure is shown in Figure 1.

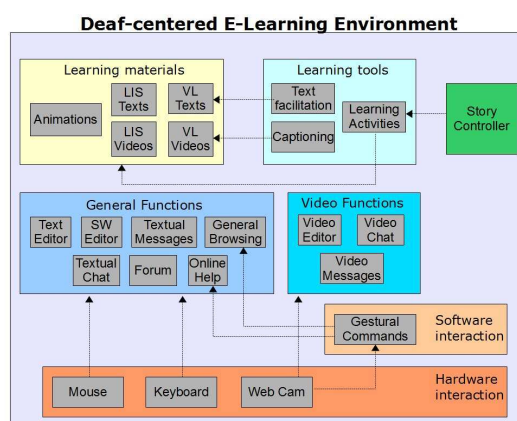


Figure 1. The DELE architecture (Bottoni et al., 2010)

We will here describe only the main features of DELE, connected to the body of theories introduced in chapter 3. From the body of theories on embodied cognition, we have picked up the suggestion to drive the learning experience back to learners’ embodied knowledge, working on storytelling and metaphors. Since DELE addresses the needs of High School and University students and young deaf professionals, the major inspiration for the navigation through the e-learning system is the University Campus metaphor, which builds the appearance of the main navigation through DELE (Figure 2).



Figure 2. Mock-up of the Campus environment in DELE

The user navigates the learning environment through an avatar, moving across the whole system under users' directions. Graphical adaptations of the metaphor in use should allow the activation of deaf people's visual-body skills, used to grasp information and to generate body-based general inferences. General image-schematic structures will be implemented (i.e. enter places, manipulate objects, meet other characters, etc.), generating an intuitive and easy interaction with the system.

Storytelling is the foundational theory of all the learning paths in DELE. The entire learning environment is made of a number of stories within stories, each implementing a different, task-specific metaphoric environment. Moving through the virtual environment, the user is asked to progress within the sub-environments (sub-stories) to accomplish several learning tasks.

These "stories" can also be developed by the courses' tutors, who can rely on a Graphical User Interface-based editor describing how these stories (i.e. learning paths) must appear. In this way, the tutor can organize the page's contents simply by connecting graphical "nodes" to each other. When the design stage is concluded, the editor automatically generates the web pages code, allowing the tutor to design a full learning path without any specific programming knowledge.

Currently, a first prototype of DELE containing a typical learning path and several learning materials has been implemented, and the "Story Controller" module allows the learning path to be dynamically generated and monitored by the system. When entering a learning story, the system shows the next possible steps in the learning process, as well as other resources (i.e. information on users who already passed these activities, help materials, in-depth activities, etc.), making the system open and flexible to the characteristics of each learner. In this way the learner is followed step by step, provided with suggestions for the subsequent module as the learner successfully completes the previous one. Progress in content acquisition and completion of complementary modules is preserved and shown as the progress within a story, visually given by a path where users need to move through, and constituting the student's portfolio.

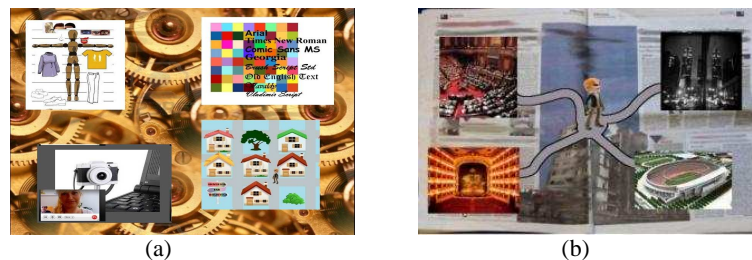


Figure 3. Two possible environments for DELE. (a) newspaper learning environment menu, (b) configuration menu

As stated above, all information will be encoded visually, hence graphical elements into the environment will have semantic relevance: background images are chosen to embed the "general feeling" of the environment task, giving additional information other than text as shown in Figure 3.

We are working on the customization of the core code of Moodle (<http://moodle.org/>) in order to setup the desired graphical appearance and logical behavior of DELE. Gesture interaction possibilities are under experimentation. Preliminary usability studies done with deaf researchers, have returned positive feedbacks. However, a more accurate usability study involving the final users is scheduled.

## 5. CONCLUSION

Literacy is a mean to access the world and its information, and, in the case of deaf people, this is made even more true by the fact that they rely especially on vision to acquire information.

The FIRB-VISEL project is exploring the effectiveness of the theoretical framework of embodied cognition and storytelling in the development of highly visual environments for deaf learners, in order to improve deaf learners' literacy skills through an opportune combination of visual information and activities on texts.

This seems to be a completely new approach in the field of deaf distant education, for which we had to face many problems for the first time, supported by very little literature on the matter. We aim this project to produce a prototype for further reflections on the use of embodiment and its entailments in education, which

seems to be a very promising approach which could be used to rethink a wide range of interactive applications, especially in the field of teaching.

## ACKNOWLEDGEMENT

The work presented into this paper is made possible by the FIRB-VISEL project (grant RBNE074T5L). We would like to thank Dr. Elena Pizzuto, who has passed away in the moment in which her efforts in building a positive and challenging working environment around the issues of deafness and literacy were beginning to bloom. We'd also like to thank all of the deaf colleagues and collaborators who have helped us in focusing our attention in the correct direction: Alessio Di Renzo, Gabriele Gianfreda, Katherine M. Groves, Luca Lamano, Tommaso Lucioi, and Paolo Rossini.

Although the authors have collaborated in editing the paper and agree on its contents, they can be considered responsible of its parts as follows: Elena Tomasuolo, introduction and chapter 2, Maria Roccaforte for chapter 3, Daniele Capuano for chapter 3 and 4, Maria T. De Monte for abstract, chapter 4, conclusion and general review of the paper.

## REFERENCES

- Bottoni, P., Capuano, D., De Marsico, M., Labella, A., Levaldi, S., 2010. Metaphor-based teaching for sensorially-critical subjects. *Proc. of ICSEC 2010, The 1st International Computer Science and Engineering Conference*. Chiang Mai, Thailand, pp.17-22.
- Caselli, M. C., Maragna, S., Volterra, V., 2006. *Linguaggio e Sordità*. Il Mulino, Bologna, Italy.
- Damasio, A., 2003. *Looking for Spinoza: Joy, Sorrow, and the Feeling Brain*. Harcourt Press, Orlando, USA.
- Fabbretti, D., Tomasuolo, E., 2006. *Scrittura e Sordità*. Carocci, Roma, Italy.
- Gärdenfors, P., 2007. Cognitive semantics and image schemas with embodied forces. In J. Krois, M. Rosengren, A. Stede, D. Westerkamp (Eds.) *Embodiment in Cognition and Culture* (pp. 57-76). John Benjamins Publishing Company, Amsterdam, Netherlands.
- Jensema, C.J., Danturthi, R.S., Burch, R., 2000. Time spent viewing captions on television programs. *American Annals of the Deaf*, 145(5), pp. 464-468
- Johnson, M., 2007. *The meaning of the body*. University of Chicago Press, Chicago, USA.
- Langer, S., 1967. *Mind: An Essay on Human Feeling*. Johns Hopkins University Press, Baltimore, USA.
- Muir, L. J., Richardson, I. E. G., 2005. Perception of sign language and its application to visual communications for deaf people. *Journal of Deaf Studies and Deaf Education*, 10(4), pp. 390-401.
- Muir, L., Richardson, I., Leaper, S., 2003. Gaze tracking and its application to video coding for sign language. *Picture Coding Symposium*, Saint Malo, France, pp. 23-25.
- Stede, A., 2007. Reading with the body. Sound, Rhythm, and Music in Gertrude Stein. In J. Krois, M. Rosengren, A. Stede, D. Westerkamp (Eds.) *Embodiment in Cognition and Culture* (pp. 143-163). John Benjamins Publishing Company, Amsterdam, Netherlands.
- Talmy, L., 1988. Force dynamics in language and cognition. *Cognitive Science*, Vol. 12, No 1, pp. 49-100.

# A PROTOCOL FOR EVALUATING MOBILE APPLICATIONS

Clare Martin, Derek Flood and Rachel Harrison  
*Oxford Brookes University*

## ABSTRACT

The ease with which mobile applications can now be created has produced an enormous number of applications from which to choose in order to accomplish any given task. It is therefore increasingly difficult for developers to establish the current state of the art when developing new software within an existing application domain, and so this paper proposes a systematic method for determining the present standard. The suggested protocol examines applications both in terms of functionality and other key attributes such as efficiency and usability, and it has been validated in two very different application domains, which are outlined here. The domains were chosen because it is believed that existing mobile applications are inadequate for them, and so the protocol has been used to determine why this is so and to identify ways in which the shortcomings can be overcome.

## KEYWORDS

Mobile Applications, Systematic Evaluation, Evaluation Protocol.

## 1. INTRODUCTION

The emergence of smart phones, with associated software developers' kits, has made it possible for a wide variety of mobile applications to be developed in a very short space of time. Mobile phone users can now complete a broad range of tasks from information retrieval to complex calculations, and this phenomenon has been summarised by one of the biggest distributors of mobile applications, Apple, in their current catchphrase "There's an app for that".

The availability of such a diversity of mobile applications can be attributed, in part, to the ease with which they can be created. Whilst the development of traditional applications requires a large investment of time and money, mobile applications can be developed quickly and with relatively little cost. Another contributory factor is the ease of distribution, facilitated by the centralised dissemination channels offered by the major phone producers, such as the App store from Apple and the Android market place from Google.

The rapid growth in the development of applications has made it increasingly difficult for developers to establish the current standard of mobile applications available for a given task; an important first step in the development of commercial applications. A systematic evaluation methodology is therefore proposed to address this issue and to allow developers to establish the present state of the art within a given application domain.

The protocol has been validated through its application to two very different domains: spreadsheets and diabetes management. Both of these form significant parts of current mobile application culture, the former being critical in the business world and the latter being crucial to the health field. Mobile spreadsheets tend to be used by business professionals who need access to numerical data whilst away from their desks, during meetings or while travelling for example. Mobile diabetes management tools are used by members of the general population who need to log healthcare information regularly, wherever they happen to be, and such people rely on the tools being very easy to use in order to self manage their condition effectively. Thus both kinds of application are vitally important to the people who use them, but little work has been done to model the usability of such tools. By applying our protocol to these contrasting application domains we can discover general issues which may arise when evaluating mobile applications, which will result in a more generic protocol.

The evaluation described here is restricted to the iOS platform used by Apple for the iPhone, iPod Touch and iPad. The reason for restricting our attention to just one platform initially was to maintain a consistent interface, and the reason for choosing this particular platform is that it currently has a significant portion of the smart phone market share, and offers a large number of existing applications in the two chosen domains. We are currently beginning to test the protocol with a range of other platforms to see how well it translates.

The methodology for evaluating mobile applications is outlined in Section 2. The two domains are then described in Section 3, together with some of the issues that exist, and Section 4 concludes this paper.

## 2. EVALUATION PROTOCOL

The steps of the evaluation protocol are as follows:

1. *Identify all potentially relevant applications.* There are a number of ways to conduct a search for appropriate applications, including a standard web search. Current software distribution methods make this easier as most of the major mobile phone platforms now have an associated online application store, such as the App store from Apple and the Android marketplace from Google.

2. *Remove light or old versions of each application.* Many software developers release trial versions of their systems, which are often free. Some of these versions include only a subset of the functionality offered by the full application whilst others allow full access to the application but for a limited time period. These types of applications should be removed.

3. *Identify the primary operating functions and exclude all applications that do not offer this functionality.* The primary operating functions include frequently used functions and also occasionally used functions that are essential for the correct operation of the system in a desired context. For example, the initial system setup might include language and currency settings that would depend upon the country of use.

4. *Identify all secondary functionality within the remaining apps.* In addition to the primary operating functions, mobile applications will offer users a range of secondary functionalities which can enhance the application. A thorough knowledge of these functions will enable the application developers to see what functionality is available and may present opportunities for missing functionality to be included in future applications.

5. *Evaluate the remaining applications in terms of:*

a. *Keystroke level modelling* (Card, Thomas et al. 1983) to estimate the time taken to complete certain tasks. This will provide a quantitative measure of efficiency of the applications.

b. *Heuristics* Standard heuristics (Nielsen and Molich 1990) can be used to evaluate desktop applications. The following heuristics, proposed by Bertini et al. (2008), can be used for the evaluation of mobile applications:

- Visibility of system status and losability/findability of the mobile device;
- Match between system and the real world;
- Consistency and mapping;
- Good ergonomics and minimalist design;
- Ease of input, screen readability and glancability;
- Flexibility, efficiency of use and personalization;
- Aesthetic, privacy and social conventions;
- Realistic error management.

## 3. CASE STUDIES

Two application domains have been selected with which to validate the above protocol. The first is the spreadsheet application, which can allow users to complete a wide variety of tasks from financial planning to statistical analysis. The second domain has much more limited primary functionality: diabetes management software allows users to manage their condition by logging daily information. The following sections outline these two domains.



### 3.1 Spreadsheets

Spreadsheets are ubiquitous software tools used for a variety of tasks from financial planning to statistical analysis. During a series of interviews (Croll 2007) with professionals within the financial sector, one participant remarked “*Put simply and succinctly, despite the higher operational risk, Excel is everywhere – it is the primary frontline tool of analysis in the financial business.*”

The mobile nature of business is increasing the need for users to access spreadsheets while on the move. Therefore mobile spreadsheet applications are becoming more important and the requirements of users are expanding to include more advanced functionality such as specialist functions and features.

An examination of this domain has been conducted for the iOS platform and identified a number of issues that should be addressed in order for mobile spreadsheet applications to reach their full potential. A full report of these findings will be published in due course.

This evaluation has found that there are many differences between mobile spreadsheet apps. Some applications allow users to view existing spreadsheets on a mobile device while others only allow users to create new spreadsheets in a mobile context. The method by which the user creates the spreadsheet also changes between apps. A Keystroke level modelling evaluation of the applications has shown that the number of keystrokes required to create a simple spreadsheet can vary by as much as 100%.

This evaluation has also identified a number of guidelines that developers should follow to improve the usability of mobile spreadsheet apps. It has been found that a number of applications do not optimise the input method for the data being entered, therefore complicating the way in which the user enters data into a spreadsheet. By considering the most common type of data to be entered into a cell, the developers could optimise the data input methods for inserting particular types of data.

The limited screen size of mobile devices has made it difficult for users to relate the section of the spreadsheet displayed on screen to the overall document. Similar problems were seen when looking at large web pages or large images (Burigat, Chittaro et al. 2008). In these domains the use of a mini-map has been proposed where a scaled down version of the document is placed in a corner of the screen. On this map the currently visible section of the document is highlighted.

The results obtained so far are limited to the iOS platform featured on mobile devices from Apple, including the iPod touch and the iPhone, as this is a commonly used platform. It is intended to do a similar study of mobile spreadsheet applications available on the Android from Google and the Blackberry.

### 3.2 Diabetes Management

Type 1 diabetes occurs when the insulin producing cells of the pancreas are destroyed leaving the body unable to control its blood glucose levels. People with type 1 diabetes have to take insulin regularly to try to stop their glucose levels from becoming too high, but if they take too much insulin their glucose levels may also drop too low, causing a number of symptoms including dizziness and palpitations.

The vast majority of patients with type 1 diabetes in the UK administer their insulin through multiple daily injections, and the remaining proportion use insulin pumps. Most people are offered a structured education programme, such as DAFNE (2011), to help them self manage their condition. This teaches them how to calculate the amount of insulin to administer at each meal according to the current blood glucose level, number of carbohydrates consumed and various other factors such as time of day, exercise and illness. The daily glucose levels are then stored in a hand-written diary which is shared with the healthcare team at regular intervals. It is surprisingly difficult for patients to keep their blood glucose levels within the target range, and yet failure to do so can lead to serious complications which are a huge burden on the health service.

Most insulin pumps come with dose calculators to help patients determine how much insulin to administer, but people on multiple daily injections do not usually have this support, and tend to do the calculations themselves. This trend is beginning to change, with the advent of glucose monitors such as the Accu-Chek Expert, manufactured by Roche, which does have a dose calculator, but its prohibitive cost has meant that it has not yet become widely used.

This need for electronic decision support, combined with the recent growth in smart phone use has led to the development of a plethora of diabetes management applications: a recent search on the App store returned 231 applications associated with diabetes, which is quite extraordinary considering the number of barriers

that have inhibited the adoption of telemedicine systems for diabetes management in the past (Belazzi 2008; Klonoff 2009). One of the key barriers has been usability, and another has been the economic implication, but the standardisation of controls and rigid human interface guidelines imposed by the providers of most phone platforms has already had a huge impact on usability, and the ease with which software applications for mobile phones can now be developed and distributed helps to mitigate against the economic factors.

The protocol described in this paper has been used to evaluate the 231 iPhone applications, and the initial filtering on primary functionality reduced the number down to 8 applications, which were further tested using KLM and heuristics. This produced some revealing differences, especially in data entry methods, and data visualisation. A number of usability problems associated with the primary tasks were also identified. Even a simple task, such as emailing data, resulted in a huge variability in the number of keystrokes, with some applications requiring a full email address to be typed in (which can be a demanding task on a phone with a small keypad) whilst others simply used a default address. The survey of secondary functionality has now been summarised in a matrix, which will be assessed by potential users in order to design an application which maximises usability whilst providing the key functionality desired by most users.

#### 4. CONCLUSIONS

This paper presents an evaluation protocol for establishing the standard of mobile applications within a particular application domain, an important first step during the creation of a new application. The ease with which mobile applications can be developed has meant that the number of applications that are available is increasing dramatically. Consequently the task of establishing the current standard of applications is becoming more difficult.

The proposed protocol is currently being applied to two application domains; spreadsheets and diabetes management. These two domains were selected as they each represent a different category of application; a general purpose application and specific application domain. By applying the protocol to both of these domains the protocol will be validated while demonstrating its flexibility.

#### REFERENCES

- Belazzi, R. (2008). Telemedicine and Diabetes Management: Current Challenges and Future Directions. *Journal of Diabetes Science and Technology* 2(1).
- Bertini, E., T. Catarci, et al. (2008). Appropriating Heuristic Evaluation Methods for Mobile Computing. *International Journal of Mobile Human Computer Interaction* 1(1).
- Burigat, S., L. Chittaro, et al. (2008). Map, diagram, and web page navigation on mobile devices: the effectiveness of zoomable user interfaces with overviews. *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services*. Amsterdam, The Netherlands, ACM.
- Card, S. K., T. P. Thomas, et al. (1983). *The Psychology of Human-Computer Interaction.*, London: Lawrence Erlbaum Associates.
- Croll, G. J. (2007). The Importance and Criticality of Spreadsheets in the City of London. *The European Spreadsheet Risks Interest Group Annual Conference*.
- DAFNE. (2011). "<http://www.dafne.uk.com/>." Retrieved 25/02/11, 2011.
- Klonoff, D. (2009). "Using Telemedicine to Improve Outcomes in Diabetes – An Emerging Technology." *Journal of Diabetes Science and Technology* 3(4).
- Nielsen, J. and R. Molich (1990). Heuristic evaluation of user interfaces. *the Conference on Human Factors in Computing Systems (CHI90)*.

# THE EFFECT OF THE DIRECTION OF A GRADIENT ON HAPTIC POINTING IN A 2D VIRTUAL ENVIRONMENT

Ludovic Potier, Olivier Gapenne, Xuan Liu and Dominique Aubert  
*Université de Technologie de Compiègne, COSTECH Laboratory  
Centre Pierre Guillaumat, 60203 Compiègne, France*

## ABSTRACT

The aim of this work was to study the role of variation in a density gradient on performances of haptic pointing in a 2D virtual environment. To carry out this study, we used the Tactos device which makes it possible to deliver tactile signals when virtual elements (for example, the points of a texture) are detected during exploration of a plane towards a target. The Tactos device also makes it possible to record the traces of the explorations. We have compared the influence of textures organized in positive or negative density gradients on performance in the pointing task. The results showed that better performances are obtained with a positive gradient.

## KEYWORDS

Pointing – Texture – Gradient – Perception – Haptic – Perceptual supplementation

## 1. INTRODUCTION

The gesture of exploring a surface with the hand associates cutaneous perception and proprioception in what has been called “haptic perception” (Overvliet, Smeets & Brenner, 2008). Almost all these studies apprehend textures by means of their roughness, which corresponds to the scale of the microstructure of the material. It has been shown that the perception of this roughness is equivalent when the subjects actively move their finger on a texture, or when a texture is applied dynamically on an immobile finger (Lederman, 1981). This result seems to show that the perceptual system involves purely a flow of stimulation produced at the level of the contact zone with the skin. However, over a century and a half ago, Weber (1834) had shown the importance of movement for the acuteness of tactile perception. His studies showed that the threshold for the detection of two distinct points decreases when passive relative movement is introduced, but decreases even more with active movement. It therefore seems premature to give no role to active movement in the perception of texture by considering that the question of dynamic touch over-rides the active/passive distinction (Hughes & Jansson, 1994).

Moreover, it is important to consider that the active exploration of a textured surface makes it possible to apprehend properties which are not limited to the microstructure (e.g. roughness), but include also the variation in density of local elements which constitutes a gradient. It is known that the use of a textural gradient can help to reproduce a position on a plane as well as the distance between two points (Schellinghouth, Smitsman & Van Galen, 1998). In addition, it has been shown that it is possible to detect the direction of a textural gradient by exploring with a single finger (Hughes & Jansson, 1994). Finally, Jin and Hughes (2005) have studied the perception of a textural gradient in one dimension. Their work shows that by controlling the speed of the relative movement of the surface with respect to the immobile hand – in order to compensate for the steepness of the gradient – subjects were still able to detect gradients, in particular positive gradients.

The aim of the present study is to determine whether the convergence of the gesture in the direction of the gradient is more effective with a positive gradient (from a low density towards a high density) or with a negative gradient (from a high density towards a low density), in order to improve navigation towards a target. The main hypothesis is that subjects will converge more rapidly on a target when the gradient is positive (with the target situated in the zone of higher density). We know that for the exploration of a gradient, there is a natural tendency to navigate in order to reveal variations in density (Schellinghouth,

Smitsman & Cox, 2005). These variations should be all the easier to detect if the texture is dense, thus in the neighbourhood of the target with a positive gradient. On the contrary, in the case of the negative gradient, the lack of information in the proximity of the target (the empty space between the points with the perceptual supplementation device produces no stimulation) should result in it taking longer to reach the target. We can also make an analogy between these textured gradients and the optic flow described in the ecological approach of Gibson (1979). This author shows indeed that the movement of an organism in an environment generates an “expansion” of its optical flow, whose density is maximal close to the focus of vision.

## 2. METHODS

### 2.1 Materials

The experiment reported here was carried out with the Tactos® device. This is a member of the technological family inaugurated by the TVSS (tactile vision substitution system) of Paul Bach-y-Rita (1972), which converts an image captured by a video camera into a “tactile image” rendered by a matrix of 400 points which can each be activated independently. Tactos offers a minimalist version of this system (16 points, identical to those used in Braille cells), which also makes it possible to explore graphic objects displayed on the screen of a computer (Figure 1).



Figure 1. Haptic exploration of a graphical object with Tactos

To use Tactos, the subject places the index finger of his/her non-dominant hand on the tactile stimulator. The other hand is used to operate the pointing system (mouse, tablet, etc.) in order to move his “receptor field” on the screen (the cursor); this triggers the activation of the tactile stimulators when black pixels are encountered. The stimulation corresponds to lifting the picots of the Braille cells. Tactos also makes it possible to record the traces of the exploration, associating at each instant the position of the receptor field and the state of the picots (the frequency of recording is 64Hz). These traces make it possible to carry out thorough analyses of the sensori-motor activity deployed by the subjects in order to perceive objects displayed on the screen (Lenay, Stewart & Gapenne, 2002).

### 2.2 Design

20 subjects age from 17 to 60 years old (17 men, 3 women, students and professors) took part in this experiment. All the subjects signed a consent formula. They were separated into two groups, one of which explored a positive gradient and the other a negative gradient.

The textures were composed of square elements with 8 pixels on each side. Three different densities were proposed, corresponding to distances between the elements of 4, 12 and 28 pixels. These densities were arranged in three circular zones. The density of the texture corresponding to each of these zones depends on the group in question (Figure 2). With the tablet employed (Intuos® size A5), the field of exploration had a diameter of 15 cm.

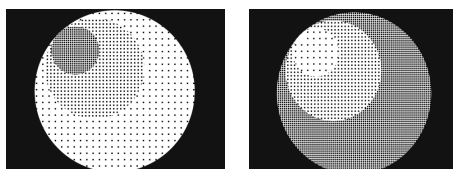


Figure 2. Examples of texture. On the left a positive gradient, on the right a negative gradient

The targets were placed in 5 different positions (Figure 3). The point of departure was always situated at the top of the field of exploration. All the participants explored each of the textures 4 times, for a total of 20 explorations. The order of the textures was randomized (5 by 4 mixed). There was a pause after each sequence of 5 trials.

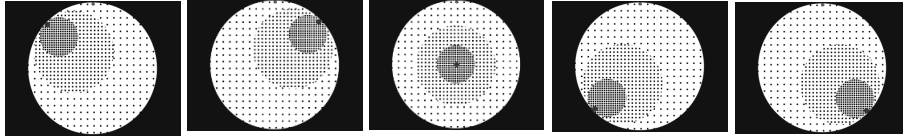


Figure 3. The five positions of the targets. From left to right: A, B, C, D and E

The experimenter welcomed the subject and invited him/her to sit down. The subject was given the instructions for the experiment to read, and was shown an example of an image such as figure 3 corresponding to the type of gradient to be explored, without indicating either the point of departure nor the target. The subject was thus informed that the textures possessed different densities, and that the target was always situated in the zone of highest density for the first group and in the zone of lowest density for the second group. The subject was also invited to keep the speed of exploration as constant as possible during the exploration.

After the subject had put on a mask, the experimenter explained three types of sounds. Three distinct types of beep signaled: the start; reaching the target; and in the event straying outside the field. The experiments started with dummy trials to familiarize the subject with the task. The subject had to succeed in reaching the target at least 2 times, in less than 2 minutes each time. If after the fifth trial the subject had still not succeeded, he was replaced by another subject (criterion of exclusion).

The recording of the traces enabled us to note the time taken to reach the target, as well as the speed of exploration in each of the zones.

### 3. RESULTS

The subjects in the first group (positive gradient) took on the average 42 seconds to reach the target, compared with 64 seconds for the second group (negative gradient). The results also showed that the time to reach the target was not the same for each of the target-positions (Figure 4). The analysis of variance showed that each of these factors had a statistically significant effect [ $F_{4,390} = 3.38$ ;  $p < 0.01$  for the positions of the targets] [ $F_{1,390} = 16.07$ ;  $p < 0.0001$  for the direction of the gradient]; there was no significant interaction between these two factors [ $F_{4,390} = 0.67$ ;  $p = 0.61$ ]. A Tuckey after-treatment showed that there was a significant difference between the target C and the target D, and between C and E [ $p < 0.05$  each time].

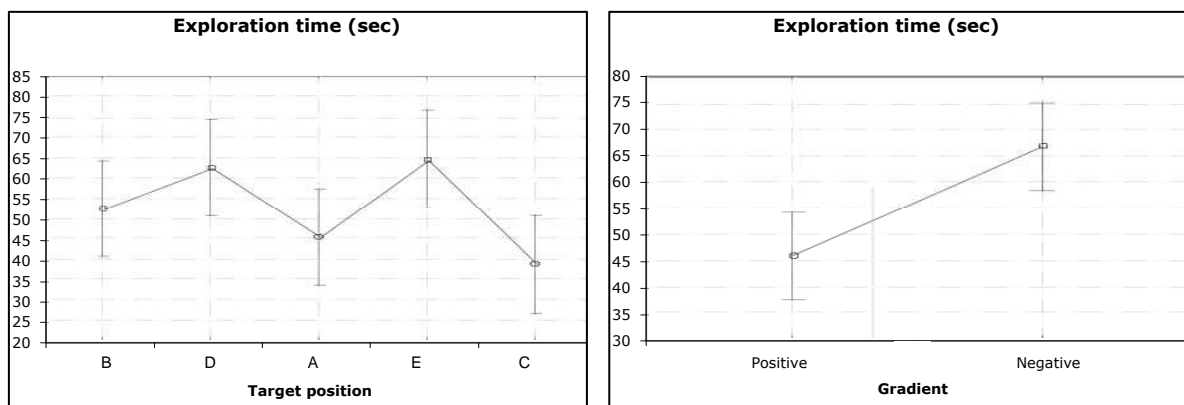


Figure 4. Influence of the conditions on the exploration times. On the left: influence of the positions of the targets. On the right: influence of the direction of the gradient

Concerning the speeds of exploration, the subjects in the first group (positive gradient) explored at a mean rate of 249 pixels/second, as against 190 pixels/second for the second group (negative gradient). The analysis

of variance showed that the mean speed was significantly less with the negative gradient [ $F_{1,58} = 4.81$  ;  $p < 0.05$ ].

In addition, we compared the speeds of exploration for each of the three zones of density. No significant effect was apparent, whatever the direction of the gradient [ $F_{2,27} = 0.45$  ;  $p = 0.64$  for the positive gradient] [ $F_{2,27} = 0.17$  ;  $p = 0.84$  for the negative gradient], which is consistent with the instructions given to maintain a constant speed of exploration.

## 4. CONCLUSION

The positive gradient has shown its greater efficiency for guiding the gesture with respect to the negative gradient. However, the exploration of textural gradients by active touch requires setting up a strategy which makes it possible to vary the tactile flux, to detect differences in density. These differences seem more evident to detect and to exploit when the texture is more dense.

The textures with a target in position C had a significant effect on the performances. That can be explained by the fact that, contrary to the other four positions where the target is placed in one of the corners, in position C the target is placed in the middle of the texture and thus has a greater chance of being encountered during the explorations.

Finally, the difference in the speed of exploration between the two groups can be explained by the initial conditions of exploration: the first group “positive gradient” started in a zone of low density, whereas the second group “negative gradient” started in a zone of high density. The speed of exploration making it possible to generate a flow of stimulations sufficient to perceive the texture was necessarily higher for the first group.

More generally, this study made it possible to show that the guidance of a gesture by textures is quite efficient in numerical 2D environments. A device such as Tactos, which has already been shown to allow persons deprived of vision to follow contours and to categorize shapes (Rovira, Gapenne, 2009), can also favour the guidance towards a surface and the discrimination of a surface by means of textures. This opens up perspectives for blind persons to have access to geometry and to geography, by optimizing the design of the “filling” of surfaces (e.g. shapes and maps).

## REFERENCES

- Bach-Y-Rita, P., 1972. *Brain mechanism in sensory substitution*, Academic Press, New York.
- Gibson, J. J., 1979. *The ecological approach to visual perception*. Editions LEA.
- Hughes, B. and Jansson, G., 1994. Texture perception via active touch. *Human Movement Science*, No. 13, pp 301-333.
- Jin, W. and Hughes, B., 2005. Effect of Density Gradients on Haptic Discrimination. *Acta Psychologica Sinica*, No. 37, pp 739-747.
- Lederman, S. J., 1981. The perception of surface roughness by active and passive touch. *Bulletin of the Psychonomic Society*, No. 18, pp 253-255.
- Lenay, C. et al., 2002. Espace d'action technique et geste perceptif. *Le geste technique : réflexions méthodologiques et anthropologiques*, Vol. 14, No. 2, pp 215-230.
- Overvliet, K. E. et al., 2008. The use of proprioception and tactile information in haptic search. *Acta Psychologica*, No. 129, pp 83-90.
- Rovira, K. and Gapenne, O., 2009. Analytical and holistic strategies employed by blind adolescents in task of tactile classification on traditional and computerized media. *Journal of Visual Impairment and Blindness*, No. 103, pp 430-435.
- Schellingerhout, R. et al., 1998. Texture information in tactual space perception. *Acta Psychologica*, No. 99, pp 93-114.
- Schellingerhout R. et al., 2005. Evolving patterns of haptic exploration in visually impaired infants. *Infant Behavior & Development*, No. 28, pp 360-388.
- Weber, E. H. (1978). *The sense of touch* [De Tactu (1834) (H. E. Ross, trans.) and Der Tastsinn (1842) (D. J. Murray, trans.)], Academic Press, New York.

# TAMULATOR: A TOOL TO MANAGE TASK MODEL-BASED USABILITY EVALUATION IN DEVELOPMENT ENVIRONMENTS

Shah Rukh Humayoun<sup>1</sup>, Yael Dubinsky<sup>2</sup>, Eli Nazarov<sup>3</sup>, Assaf Israel<sup>3</sup> and Tiziana Catarci<sup>1</sup>

<sup>1</sup>*Dipartimento di Informatica e Sistemistica "A. Ruberti",*

*SAPIENZA - Università di Roma, Via Ariosto – 25, 00185, Roma, Italy*

<sup>2</sup>*IBM Research – Haifa, Mount Carmel, Haifa 31905, Israel*

<sup>3</sup>*Computer Science Department, Technion – Israel Institute of Technology, Haifa 32000, Israel*

## ABSTRACT

We present the TaMULATOR tool for managing and automating an end-to-end task model-based usability evaluation at the Integrated Development Environment (IDE) level through providing a set of APIs and interfaces for tagging tasks and variables, creating task models written in a formal task modeling language TaMoGolog, defining evaluation experiments, recording experiment data, and automatically analyzing the results. We found that using TaMULATOR tool by development teams enables ongoing professional usability evaluation of the developed products.

## KEYWORDS

Usability evaluation, task models, TaMoGolog, TaMULATOR.

## 1. INTRODUCTION

High-level usability is acknowledged as a significant feature of software products. Normally, usability evaluation is performed using existing rigorous approaches and techniques that enable the process of defining and running experiments, collecting and analyzing results, and making decisions regarding which feedback to adopt and to what extent [4]. However, in many cases, evaluating usability is performed manually [7] and due to budget and schedule concerns, it is sometimes neglected. Automating evaluation techniques throughout the development lifecycle provides several benefits, e.g.; reduced development costs and time, improved errors tracing, better feedback, and increased coverage of evaluated features [7].

Our research approach [5, 6] involves defining evaluation experiments and running them from within the Integrated Development Environment (IDE) that equips the development team with the mechanism to monitor and control a continuous evaluation process tightly couple with the development process, thus receiving on-going user feedback while continuing development. To fill the gap of formal modeling user tasks in experiments, we propose a tool, called TaMULATOR, that works at the IDE level to automatic collecting and analyzing users' and system activities and behavior through task model-based evaluation approach for finding usability flaws and errors more accurately. The automatic analysis of finding usability issues enables to draw conclusions to derive relevant development tasks for further improvements in the developing product more efficiently and effectively. Note that we use the term *usability evaluation* for the evaluation of both product usability and functionality. We use experiments to find usability issues and serve as a kind of acceptance test for the developed features.

There are only a handful of tools available that support and automate the process of usability evaluation through task models. USINE [8] takes input the task models and the logs generated by users, and then the designer creates a log-task table. USINE precondition table emphasizes the possibility of doing one task before others. RemUSINE (Remote USINE) [10] is an extension of USINE and provides the support of remote usability evaluation. MultiDevice RemUSINE is a tool [11] for mobile applications that includes the possibility of detecting those environment conditions that could affect users' interaction with mobile applications. AWUSA [1] focuses on remote task-based usability evaluation and targets websites rather than

system applications. ReModel [1] performs remote usability evaluation via a client-server architecture where the server contains task models and the targeted task models are delivered to the client via a corresponding graphical user interface.

The main differences from previous tools are that TaMULator works at the IDE level to automate and manage the complete end-to-end evaluation life-cycle, defines and tags tasks and variables at the code level, uses precondition axioms (a strong form of constraints) of tasks and postcondition effects on variables for automatic analyzing users' behaviors and system functionalities more accurately.

In Section 2, we briefly describe the TaMoGolog language for writing formal task models to be used for writing user tasks for evaluation experiments. In Section 3, we present TaMULator tool and show its working to managing and automating task model-based evaluation at the IDE level. In Section 4, we briefly provide the evaluation of the TaMULator. We conclude in Section 5.

## 2. TAMOGOLOG TASK MODELING LANGUAGE

We provide a formal task modeling language **TaMoGolog (Task Modeling Golog)**, built on top of the Golog-family [2, 3, 9, 12] of high-level programming languages. TaMoGolog provides well-defined syntax and semantics, enables precondition axioms of tasks, states postcondition effects to variables due to the completion of tasks, and provides a rich set of operators. TaMoGolog distinguishes tasks into three main categories: *unit* tasks, denoted as  $\mu$ , are considered to be performed in an atomic manner; *waiting* task, denoted as  $\varpi$ , wait either for a particular event to happen or for some set of conditions to be fulfilled; while *composite* tasks, denoted as  $\Gamma$ , handle the structural behavior of the task model. TaMoGolog provides a rich set of operators, mostly obtained from the Golog-family, useful for constructing complex task models. The TaMoGolog set of operators consist of: waiting/testing condition  $\langle \varphi \rangle$ , sequence  $[\langle G1 \rangle; \langle G2 \rangle]$ , internal nondeterministic choice  $[\langle G1 \rangle \mid \langle G2 \rangle]$ , external nondeterministic choice  $[agt \langle G1 \rangle \mid \langle G2 \rangle]$ , conditional choice  $[\text{if } \phi \text{ then } \langle G1 \rangle \text{ else } \langle G2 \rangle]$ , internal nondeterministic choice of arguments  $[\pi x. \Gamma(x)]$ , external nondeterministic choice of arguments  $[agt \pi x. \Gamma(x)]$ , internal iteration  $[\Gamma]^*$ , external iteration  $[agt \Gamma]^*$ , conditional loop  $[\text{while } \phi \text{ do } \Gamma]$ , concurrency  $[\langle G1 \rangle \parallel \langle G2 \rangle]$ , concurrency with priority  $[\langle G1 \rangle \parallel (\langle G2 \rangle)]$ , concurrent iteration  $[\langle G1 \rangle^* \langle G2 \rangle]$ , external priority concurrency  $[agt \langle G1 \rangle \parallel (\langle G2 \rangle)]$ , external concurrent iteration  $[agt \langle G1 \rangle^* \langle G2 \rangle]$ , interrupt  $\langle \phi \rightarrow \langle G \rangle$ , and failure handling  $[\langle G1 \rangle \triangleright \langle G2 \rangle]$ .

In brief, a task model in TaMoGolog consists of the following predicates and definitions: name of the *task model*; sets of *unit*, *waiting*, and *composite* tasks; definition of each *waiting* and *composite* task through Golog procedure definition [9]; set of *precondition axioms* for each unit task; *postcondition effects* on variables (relational and functional fluents) for each unit task; external entities and their responsible tasks definitions; initial values of variables; any optional domain knowledge representation; task model structure and relationship between tasks through above set of operators.

TaMoGolog semantics are also based on Golog-family [2, 9], where a *unit* task executes if all precondition axioms are true and then effects are shown on related variables (fluents) after execution. Unit tasks are performed in atomic manner while composite tasks are performed in step-by-step execution.

## 3. THE TAMULATOR TOOL

TaMULator is a Java-based tool that works at the development environment level to manage and automate a task model-based usability evaluation process. TaMULator manages an end-to-end usability evaluation process through five phases named as: *tag*, *task model creation*, *evaluation-experiment creation*, *run & record*, and *analysis*. TaMULator consists of five main modules: *TaMoGolog-Compiler*, *Task-Model*, *Experiment*, *Recorder*, and *Analyzer*. Along the interfaces provided by these modules, TaMULator also provides APIs of a Java library that is used at the source code level to help the development team to perform different things; such as tagging tasks and variables; and to integrate the evaluation process tightly in the development environment. The detail of TaMULator work in each phase follows:



**Phase 1 - Tag:** TaMULATOR allows the development team to tag unit tasks and variables (possible candidates to be used in precondition axioms of tasks and postcondition effects to variables) of interest. It provides an API, called **TTag-API (Task Tagging API)**. Developers can use it to tag unit tasks and variables by simply stating their name or can use AspectJ (refer to <http://www.eclipse.org/aspectj/>) facility, a Java extension, to define multiple *pointcuts* and *advices* that inject small snippets to let TaMULATOR know that an event has occurred. To perform tagging, developers need to “wire” their programs in “interesting” locations (for tasks and variables) throughout their source code. The TaMULATOR tool leaves it to the software team to define those points of interest (e.g., method invocations, object state changes, or certain events). Figure 2 (left) shows a screenshot of the TaMULATOR demonstrating a typical AspectJ hook.

This tagging facility can be used at any stage of the development process, and does not require any internal intervention in the program code. Only the tagged set of tasks and variables are later available for using in task models. Figure 1 (left) shows a TaMULATOR screenshot of available tagged set of tasks and variables.

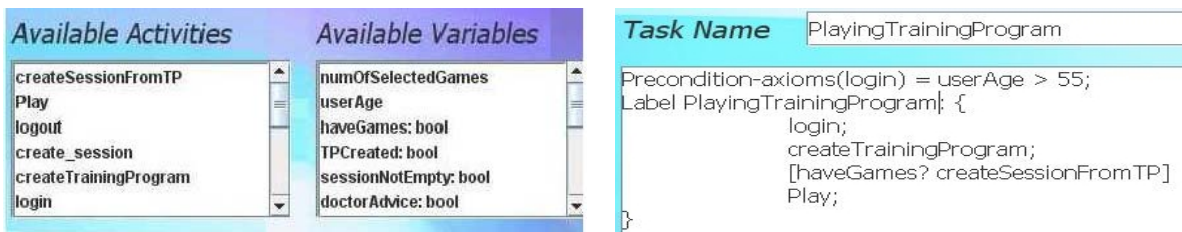


Figure 1. (left) tagged set of tasks (activities) and variables (right) A task model written in TaMULATOR

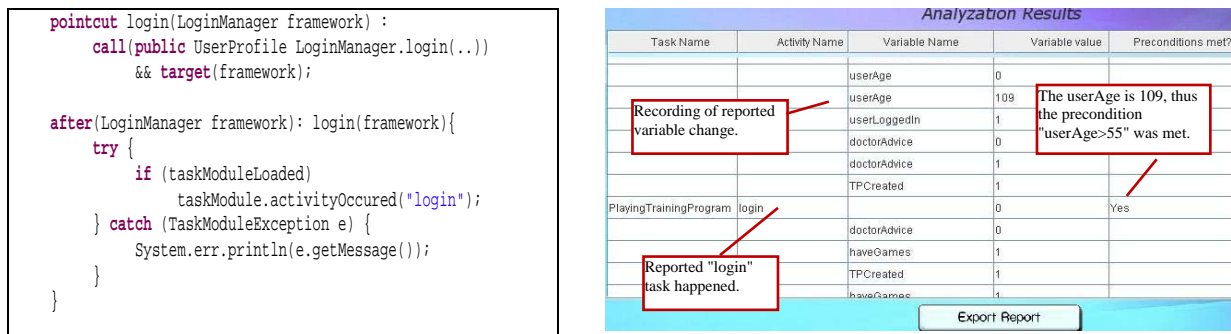


Figure 2. (left) Screenshot of a typical AspectJ hook (right) Analyzer output of an experiment

**Phase 2 – Task Model Creation:** In this phase, the software team creates task models, written in TaMoGolog, for evaluation experiments. Each evaluation experiment contains one or more scenarios the software team wishes to track, where each scenario is reflected by a task model created in this phase. The process of defining task models in TaMULATOR is very easy; simply write a TaMoGolog-based script that defines the task-model structure with temporal relations in tasks using TaMoGolog set of operators, precondition axioms, and postcondition effects and let TaMULATOR compile it. The *TaMoGolog-Compiler* module compiles the TaMoGolog-based script of a task model that can be understood by other modules. Figure 1 (right) shows a TaMULATOR screenshot of a simple task model written in TaMoGolog.

The *Task-Model* module keeps the compiled task models and provides to developers the opportunity to associate the compiled task models to evaluation experiments. TaMULATOR provides an API, called **TM-API (Task Model API)**, which developers can use to change compiled task model parameters (e.g., *task's* name) and can export the task models structures as a tree using the Java *JTree*<sup>1</sup> class.

**Phase 3 – Evaluation-Experiment Creation:** The TaMULATOR provides the facility to create evaluation experiments so that evaluating users can perform different tasks on the target application, while the TaMULATOR records user and system activities and behavior for finding out usability issues. Each evaluation experiment is associated with one or more task models from the *Task-Model* module, and is independently responsible for managing its own task models. The evaluating users are supposed to perform tasks on the

<sup>1</sup><http://download.oracle.com/javase/1.4.2/docs/api/javaw/swing/JTree.html>

evaluating application as described by the scenarios in the attached task models. TaMULATOR provides another API, called **EExp-API (Evaluation-Experiment API)**, for managing evaluation experiments and associated task models. The *Experiment* module is responsible for managing these evaluation experiments.

**Phase 4 – Run & Record:** The TaMULATOR tool also manages execution of the evaluation experiments by evaluating users, where they perform different tasks on the evaluating application to achieve the desired goal(s) of each attached task model. During running of these experiments, each tagged task that was enabled by the user or by the system, and the values of variables related to precondition axioms and postcondition effects (before executing the task and after completing the task) are reported and recorded by TaMULATOR. The *Recorder* module is responsible for recording all this data. Currently, any task record comprises the following 4-tuple: (name, timestamp, precondition axioms' status, task-model set<sup>2</sup>), while variable records for checking postconditions comprise a 3-tuple: (name, new value, timestamp).

**Phase 5 – Analysis:** During this phase, TaMULATOR takes the recorded data of the evaluation experiment and gives the results after analyzing the recorded data with original task models attached to that experiment. TaMULATOR provides a built-in *Analyzer* for this automatic analysis. The recorded data can also be exported into a CSV format for manual analysis. At any time, the development team/evaluator can issue an analysis of the recorded data against any task model/experiment. The *Analyzer* compares and analyzes the structure of the task model to the recorded data. Thus, the *Analyzer* checks whether the recorded data is consistent with the *task structure*, by checking that the appearance of the recorded unit tasks are in the same order as in the task structure, making sure that precondition axioms were met and postcondition variables possess the desired values. If any of these conditions are not met, the scenario for the task model is considered as not properly fulfilled, compared to what the designer/evaluator wanted. Figure 2 (*right*) shows a TaMULATOR screenshot of *Analyzer* output. This automatic analysis process helps the development team/evaluator to find usability issues in the targeted application, and is especially useful for locating the points where users normally make mistakes, so the development team can take care of those places in the forthcoming development iterations.

## 4. EVALUATION

In a case study, six development teams<sup>3</sup> developed their software project through agile development approach while using TaMULATOR to evaluate their developed project. The usability evaluation of the product that is being developed was implemented as part of this project. Following are the main practices we used: 1) Iterative design activities that include cycles of development that contain development tasks that were derived from usability evaluation. 2) Role holders in the subject of usability evaluation and using TaMULATOR. 3) Measurements that were taken by the role holders as part of fulfilling their responsibilities.

In addition to the developing their project, named 'Brain Fitness-Room' aims to develop a system for maintaining and strengthening memory and brain capabilities as well as identifying any decline in these capabilities, the subject of usability evaluation using task models was presented to the teams. In the first development iteration, all teams had to develop a tool with a Java library to enable writing basic task models in a formal way. Based on the teams' work, one tool was selected (TaMULATOR) and all teams used this tool during two iterations of evaluation.

In the final retrospective on the course, team members were asked to grade their satisfaction between 1 to 5 (very satisfied) with respect to the project topic, course methodology (agile, time management and early detection of problems, emphasis on testing and usability), tools that were used (Trac and Moodle), and the services in the physical lab they worked in. 32 team members answered and the average grade for the methodology was high (4.09) (for project topic 4.36, tools 3.98/3.28 respectively, and for lab services 2.66). Specifically, regarding the roles that concerned with usability, team members referred to the importance of learning and dealing with usability while developing. Following are some of their comments on this matter: "It is important to get feedback from the users...", "It does not matter how good the product is, [people] will use it only if it is simple and user friendly. A lot of things that seem clear to developers are not clear to the

<sup>2</sup> All task models that have the mentioned activity.

<sup>3</sup> The team members were 4<sup>th</sup> year CS-major students participating in the 'annual project in software engineering' course of the Computer Science Department at Technion, IIT.

end users,” and “The role of being in charge of the evaluation experiment was an important role with which we specified the usage of our system by the user.”

## 5. CONCLUSION

In this paper, we presented TaMULATOR tool that manages and automates an end-to-end task model-based usability evaluation process at the IDE level, and provides an automatic analysis of the users’ and system activities and behavior to find out usability flaws and errors more accurately and effectively. The features that distinguish TaMULATOR from other tools include providing APIs to work at the source code level, writing task models through a well-defined formal language (TaMoGolog), and the usage of precondition axioms and postcondition effects on variables during the analysis phase to highlight usability issues and to check product functionality.

In future, we intend to continue work to provide further functionalities; such as; fully support of TaMoGolog, creation of task models in visual mode, improvement in the *Analyzer* module for giving suggestions based on the analyzed results.

## REFERENCES

- [1] Buchholz, G., Engel, J., Martin, C., and Propp, S. 2007. Model-based usability evaluation: evaluation of tool support. Proc. of the 12<sup>th</sup> int. Conf. on Human-Computer interaction: interaction Design and Usability, Beijing, China, July 22 - 27, 2007.
- [2] de Giacomo, G., Lespérance, Y., and Levesque, H. J. 2000. ConGolog, a concurrent programming language based on the situation calculus. *Artif. Intell.* 121, 1-2 (Aug. 2000).
- [3] de Giacomo, G., Lespérance, Y., and Pearce, A. R. 2010. Situation Calculus Based Programs for Representing and Reasoning about Game Structures. KR 2010.
- [4] Dix, A., Finlay, J.E., Abowd, G.D., and Beale, R. 2003. Human Computer Interaction, 3<sup>rd</sup> Edition, Prentice Hall.
- [5] Dubinsky, Y., Catarci, T., Humayoun, S. R., and Kimani, S. 2007. Integrating user evaluation into software development environments, 2nd DELOS Conference on Digital Libraries, Pisa, Italy.
- [6] Humayoun, S. R., Dubinsky, Y., and Catarci, T. 2009. UEMan: A tool to manage user evaluation in development environments. ICSE’09, May 16 - 24, 2009.
- [7] Ivory, M. and Hearst, M. 2001. The State of the Art in Automating Usability Evaluation of User Interfaces. *ACM Computing Surveys*, 33, 4, pp. 470-516, Dec. 2001.
- [8] Lecerof, A. and Paternò, F. 1998. Automatic Support for Usability Evaluation. *IEEE Trans. Softw. Eng.* 24, 10 (Oct. 1998), pp. 863-888.
- [9] Levesque, H. J., Reiter, R., Lespérance, Y., Lin, F., and Scherl, R. B. 1997. GOLOG: A Logic Programming Language for Dynamic Domains. *Journal of Logic Programming*, vol. 33.
- [10] Paternò, F. and Ballardin, G. 2000. RemUSINE: a bridge between empirical and model-based evaluation when evaluators and users are distant. *Interacting with Computers*, Vol. 13, Iss. 2, Dec. 2000, pp. 229-251.
- [11] Paternò, F., Russino, A., and Santoro, C. 2007. Remote evaluation of mobile applications. In Proceedings of the 6th international Conference on Task Models and Diagrams for User Interface Design, Toulouse, France, Nov. 07-09, 2007.
- [12] Sardina, S., and Lespérance, Y. 2009. Golog Speaks the BDI Language. Proc of PROMAS’09.

# ON THE FACTORS OF THE WEBPAGE NAVIGATIONAL PERFORMANCE

Wen-Te Chang<sup>1</sup>, Kuo-Chen Huang<sup>1</sup> and Chan-Li Lin<sup>2</sup>

<sup>1</sup>Product Design dept. Ming Chuan University,

Wen-Te Chang, No. 17, Lane 182, Sec. 1 Chung-San Rd., Sec 1, Shu-Lin, 238, Taipei, Taiwan ROC

<sup>2</sup>Digital Multimedia Design dept., China University of Technology, Taiwan ROC

## ABSTRACT

This study addressed the relationship between gender and topic preference in terms of webpage navigational performance. The authors investigate the correlation between topic preference and gender through questionnaire survey, and test their influences on webpage navigational performance by webpage wayfinding tasks. As the subjective questionnaire test result indicated in this study, males favored but females disliked the vehicles topic, whereas females favored but males disliked the cosmology topic, and both group (female and male) all favored tourism topic. Through the three topics, vehicles, cosmology, and tourism topic, web page navigation tests, the significant interaction between gender and topic factors was found. It was not neglected either that the difference between gender and vehicle topic was significant, but the difference on the tourism and cosmology was not significant. It was also noted from MANOVA analysis that topic preference has impact on the male group webpage navigational performance but not on the female group. Male group tend to be affected by the preference of topic thus have better performance on the tourism and vehicle topic and worst performance on the cosmology topic.

## KEYWORDS

Web, Navigation, Topic Preference, Topology, Gender.

## 1. INTRODUCTION

This study addressed the relationship between gender and topic preference and the interaction through webpage navigation tests and subjective questionnaire survey.

## 2. BODY OF PAPER

### 2.1 Materials, Design, and Data Analysis

A mixed factor (2x3) experimental design was adopted in this study. The independent variable was gender and topic (Tourism, Cosmology, and Vehicles). The dependent variables were the mean time of locating specific nodes (Shin, et al., 1994). On the one hand, between subject design was adopted for the variables of gender. On the other hand, within subject design was used for the variables of topic.

Eighty-nine participants (44 females and 45 males) were recruited for the experiment. All were college students from the design school of Ming Chuan University, Taiwan, who spent an average of 5.34 hours (SD=1.98) in daily computer and Internet browsing. They ranged in age from 19 to 23 years (M=20.84, SD=1.02). All participants were paid (150 New Taiwanese dollars/ \$5 per hour) for their participation.

#### 2.1.1 Topic Preferences

As previous studies have indicated (Salmerón et al., 2006) topic preference affects navigational performance significantly, as does gender (Coluccia, & Louse, 2004). To investigate this relationship, six experts (two

each from the fields of psychology, human factors, and interface design) were invited to discuss the most popular topics on Internet search engines, such as Yahoo, Yam, and Google. Seven topics were selected: Politics, International News, Economics, Cosmology, Entertainment, Vehicles, and Tourism. Two hundred and ninety-three participants (148 females and 145 males) were recruited for a preliminary questionnaire investigation. The participants were asked to evaluate the topics from most to least favored. All of the participants were college students from the design school of Ming Chuan University who spent 5.19 hours (SD=1.14) daily on average using a computer. Their ages ranged from 19 to 25 years (M=21.3, SD=1.15). Figure 1 shows the value and correlation of the topic and gender factors. The questionnaire was coded on seven scales, with the favorite topic receiving 7 points and the least favorite 1 point.

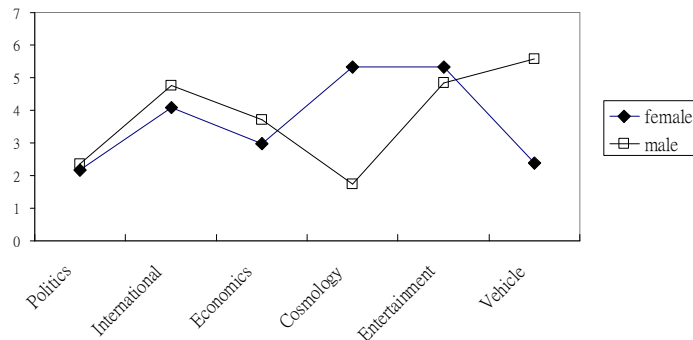


Figure 1. Values of The Topic and Gender Factors (level: 1~7)

The questionnaire results were analyzed at an  $\alpha$  level of 0.05. The significant data were analyzed to determine the topic most favored by females that males disliked, the topic most favored by males that females disliked, and the favorite topic of both males and females, as well as selected topics for further experimental design and development. The seven topics showed a main effect,  $F(6, 1746)=174.96, P<.05$ . The interaction between topic and gender was also significant,  $F(6, 1746)=142.8, P<.05$  for vehicles and cosmology. Males favored and females disliked the vehicles topic, whereas females favored and males disliked the cosmology topic. Favorable levels for both politics and economics were lower than the mean ( $<4$ ), whereas those for international news, entertainment, and tourism rated higher ( $>4$ ), with tourism the highest ( $>5$ ). Based on the data analysis, the topics of tourism, cosmology, and vehicles were chosen for further Web design.

### 2.1.2 Topological Structures

Before constructing the webpage, the authors need to define the most appropriate number of hyperlink nodes. Based on the previous studies (McDonald, & Stevenson, 1998) and the advice of the invited experts (six experts, two each from the fields of psychology, human factors, and interface design), the authors finally chose 60 nodes for the experimental tool design.

The main text content was collected from Internet searches, with most text in Chinese but with a few English words used due to the type of serial or original content. Each Web page represented one node in the hyperlink system. Each page had between 400 and 447 Chinese Ming characters (M=421; SD=14) in 14-point type. The total Web content amounted to 25,267 words.

The web pages were created using Hypertext Markup Language (HTML) based on hierarchical topology. The final experimental materials were approved by the same group of six experts, who met after iterative design proposals and pilot tests.

The topological web page design featured each node in a fixed hierarchical sequence. Each individual Web page could be accessed by clicking on "forward" and "back" buttons at the bottom of the pages, which caused the next or previous page to be displayed. The hierarchical version links were based on one parent node for every two child nodes, resulting in 20 nodes with 19 links (Figure 2) for each topic. The three topics, with 20 links each, had a total of 59 extended hyperlinks.

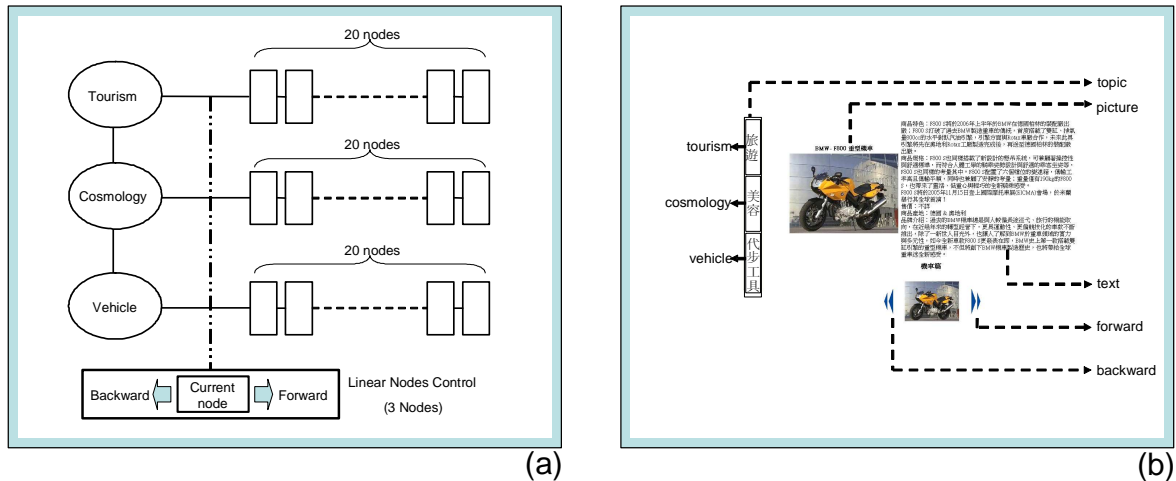


Figure 2. Illustration of the web page design: (a) the topological structure; (b) the interface and layout

All documents were run on Microsoft’s Internet Explorer 6.0 browser and were displayed on a 17-in LCD color monitor. The resolution was 1024x768 pixels at a frame rate of 85Hz. The activities of participants were monitored throughout the experiment.

The nature of the experiment and task specifications was first briefly explained. Then, the participants were asked to learn to interact with the user interface by conducting practice tasks. No time constraint was imposed during the warm-up session. The session ended when the participant felt confident of having a basic understanding of the interface hierarchy and relationships among the various nodes. After the participants had become familiar with the Web design, the experiment formally began. Each task was explained on a paper sheet, and the participants followed the list order to complete the navigation tasks. Participants were required to complete all of the interaction tasks. The full experiment lasted approximately 1hr for each of the participant.

## 2.2 Results and Analysis

Table 1 presents the means and standard deviation of navigation time by split data of gender and topic. Table 2 further presents the mean difference of navigation time by gender and topic factors.

Table 1. Means (Standard Deviation) of navigation time by split data

Gender	N	Tourism	Cosmology	Vehicle
Female	44	20.43 (6.03)	25.52 (8.28)	27.76 (9.89)
Male	45	18.36 (9.53)	26.16 (11.68)	17.08 (5.62)

(unit: second)

Table 2. Mean difference of post hoc test by split data

Gender	N	Mean Difference			
		Topics	Tourism	Cosmology	Vehicle
Female	44	Tourism	-	-	-
		Cosmology	-5.091(*)	-	-
		Vehicle	-7.326(*)	-2.235	-
Male	45	Tourism	-	-	-
		Cosmology	-7.800(*)	-	-
		Vehicle	1.274	9.074(*)	-

(unit: second)

Based on estimated marginal means, adjustment for multiple comparisons: Bonferroni.

\* The mean difference is significant at the .05 level.

## 2.3 Discussion

Previous studies have suggested that males have better spatial strategy shifting capacity (Scali, Brownlow, & Hicks, 2000) and visual working memory (Coluccia, Bosco, & Brandimonte, 2007) than females. Nevertheless, related issues such as topic preferences of the webpage information based on the gender difference was neglected by most researches.

It was demonstrated from the subjective questionnaire test result in this study, that the topics preference of the males and females were contradictory on cosmology and vehicle topics (female disliked the vehicles topic, but males disliked the cosmology topic), whereas both female and male all favored tourism topic. Through the three topics, vehicles, cosmology, and tourism topic, webpage navigation tests, the significant interaction between gender and topic factors was found. Figure 3 shows the interactive effect between gender and topic factors. It was not neglected either that the difference between gender and vehicle topic was significant, but the difference on the tourism and cosmology was not significant. Previous studies argued that individual topic preference significantly affects navigational performance (Hidi, 2001; Salmerón, Kintsch, & Cañas, 2006).

To further explore the correlation between the gender and topic preference effect, it was noted from MANOVA analysis that topic preference has impact on the male group webpage navigational performance but not on the female group. Male group tend to be affected by the preference of topic thus have better performance on the tourism and vehicle topic and worst performance on the cosmology topic navigation test. This is consistent with the argument of the previous studies that topic interest generates spontaneous attention resulting in faster information processing (shorter searching time needed) (Hidi, 2000, McDaniel, Waddill, Finstad, & Gourg, 2000).

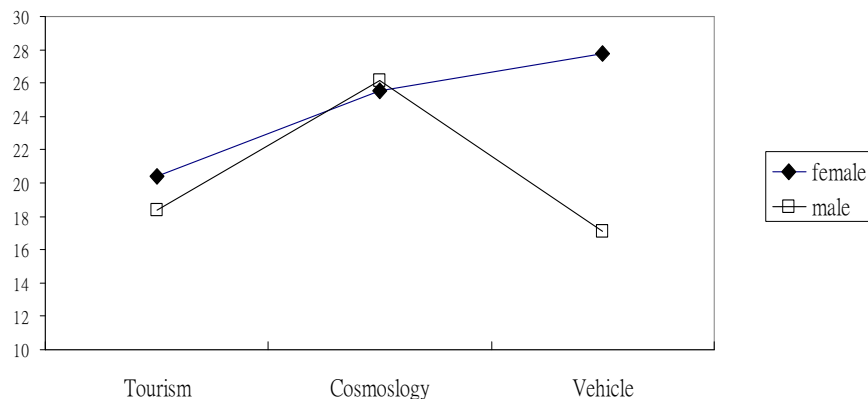


Figure 3. The correlation between gender and topic factors. (unit: second)

## 3. CONCLUSION

From the webpage design practice perspective, to compare the gender difference on webpage navigational performance without the consideration of other related user-centered factor will be problematic. Because new webpage interfaces design involved not only topological structure planning issue, but also environmental and individual differences. To integrate the objective experimental research method and more subjective realistic factors (e.g., topic preference or wayfinding strategy) can then ensure the better holistic design outcome.

## ACKNOWLEDGEMENT

Financial support of this research study by National Science Council under the grant NSC 99-2410-H-130-042- is gratefully acknowledged.

## REFERENCES

1. Coluccia, E., & Losue, G., 2004. Gender differences in spatial orientation: A review. *Journal of Environmental Psychology*, 24(3), 329-340.
2. Coluccia, E., Bosco, A., & Brandimonte, M. A., 2007. The role of visuo-spatial working memory in map drawing. *Psychological Research*, 71, 359-372.
3. Hidi, S., 2000. An interest researcher's perspective on the effects of extrinsic and intrinsic factors on motivation. In Sansone, C., & Harackiewicz, J. M.,(eds.), *Intrinsic Motivation: Controversies and New Directions*, New York: Academic Press, 309-339.
4. Hidi, S., 2001. Interest, reading, & learning: theoretical and practical considerations. *Educational Psychology Review*, 13, 191-209.
5. McDonald, S., & Stevenson, R. J., 1998. Effects of text structures and prior knowledge of the learner on navigation in hypertext. *Human Factors*, 40(1), 8-27.
6. Salmerón, L., Kintsch, W., & Cañas, J. J., 2006. Reading strategies and prior knowledge in learning from hypertext. *Memory & Cognition*, 34(5), 1157-71.
7. Scali, R. M., Brownlow, S., & Hicks, J. L., 2000. Gender differences in spatial navigational performance as a function of speed or accuracy orientation. *Gender Roles*, 45, 359-376.





# **Reflection Papers**



# SENSOR-BASED ESTIMATION OF PSYCHOLOGICAL STATES

Javier Eguez Guevara and Hiroyuki Umemuro

*Department of Industrial Engineering and Management Tokyo Institute of Technology  
2-12-1 Ookayama, Meguro-ku, Tokyo, 152-8550, JAPAN*

## ABSTRACT

The purpose of this paper is to draw attention to the importance of sensory technology for the purpose of estimation of psychological states such as affect, personality, and mental health. Firstly this paper summarizes conventional methods used to measure psychological states, and describes related known issues. Furthermore, referring to recent studies that tried to incorporate sensory technology to estimate psychological states, it discusses the advantages and disadvantages of sensor-based estimation methods, in comparison with conventional ones.

## KEYWORDS

Psychological states estimation, sensory technology, human behavior, affect, personality, mental health

## 1. INTRODUCTION

Techniques that attempt to classify and measures psychological states require acknowledgement of its dynamic behavior, and therefore it is still necessary to propose less obtrusive methods which achieve results of greater accuracy. Although both experimental and non-experimental methods have been extensively used in the process of measuring behavior in the last century, these methods are generally invasive, and require huge amount of time and resources for its operation. Despite the countless number of characteristics defining the human being, this paper focused on three behavior-based psychological states, namely, affect, personality, and mental health.

On the other hand, there are a variety of possibilities which utilize sensors for the recognition of psychological states. Occupancy, movement and orientation, touch, position, identity, affect, and others, are contexts where recognition mechanisms such as walking gait patterns, trunk movement, gesture analysis and others utilize sensors in different ways. Although a number of issues have been found in the utilization of sensory technology for psychological states estimation (Wilson, 2008), its adequate utilization has facilitated the measurement of such states providing unobtrusive and time-efficient mechanisms to estimate psychological states through the measurement of a selected set of body movements and behaviors.

The purpose of this paper was to draw attention to the importance of the incorporation of sensory technology in the estimation of psychological states by outlining limitations, concerns, advantages and disadvantages of methodologies in recent studies.

## 2. CONVENTIONAL ESTIMATION METHODS

### 2.1 Affect Estimation

There are several affect measuring techniques, where the brain is perhaps the most fundamental source of emotion (Brave & Nass, 2008). Another mechanisms used for the study of emotion is the measurement of the autonomic nervous system, the use of facial expression, voice, and body movements. Picard (1997) referred to the complexity related to the physical aspects of the estimation of human behavior, arguing that emotion does not map to a fixed form of physical communicating means. In general, emotion and other psychological

states are not always displayed in consistent ways, which complicates the process of behavior recognition. Some of the complicating factors are the intensity of the emotion; the type of emotion which may emphasize on an specific characteristic of a given emotion; the way the state was induced; and the social display rules by which a person express or suppress emotion (Picard, 1997).

Also, the characteristics of the scenarios inside and outside the laboratory settings pose different challenges. Most of the behavior-related studies have been confined to artificial lab scenarios with important limitations (Picard, 1997). Many emotion recognition experiments done in laboratory settings resort to subjects whose emotions intensity is strengthened or controlled, and therefore can be repeated at convenience of the research goal. Subjects in laboratory settings might exhibit a much smaller repertoire of emotions compared to those they would express in their natural world; or they might express emotions they think they should express, instead of letting them arise naturally. These issues were defined as input-specific, namely, display rules, deception, and systematic ambiguity (Fragopanagos & Taylor, 2005).

The ideal study should be conducted by means of real-life observation, which many theorists have deemed impossible (Wallbot & Sherer, 1989). Picard (1997) added that wearable computers offer the possibility of collecting data from people as they engage in natural and social interactions. However, it is of critical importance to note that wearable devices should avoid annoying or invading their privacy. One mechanism analyzed to prevent this kind of implications is to avoid computers from storing information, and rather resort to real-time affective analysis, which is an area of current research (Foner, 1996; Picard, 1997). Wearable devices may as well bring undesired situations, and therefore it is necessary for users to be made aware of both the benefits and disadvantages of these technologies.

## 2.2 Personality and Mental Health Estimation

Self report inventories, which have been also called objective tests (Meyer & Kurtz, 2006), have been firmly established as the preferred personality (Winter & Barenbaum, 1999), and mental health measuring methods. However, specific issues regarding these methods have been pointed out. Meehl (1945) contends that structured personality tests lack of a restriction-free nature, which can not accurately describe the examinee behavior. This would falsify the actual relationship between what a man says and what he is. This adds to arguments relating limitations in self-knowledge or self-perception, and personal dynamics.

Empirical evidence suggests that an individual's behavior is not constant from situation to situation. Mischel (1973) argued that there is evidence supporting substantial changes in practically all of the dispositional measures of personality of characteristics of individuals over time and across situations.

Also John and Srivastava (1999) argued that the number of scales designed to measure personality has escalated without an end in sight. While researchers have experienced little guidance and a lack of an overall rationale, it is noted that a systematic accumulation of findings is necessary in order to cope with the difficulty related to the bewildering number of concepts and scales. This is aggravated by the fact that often scales under the same name measure concepts representing different meanings (John & Srivastava, 1999).

Likewise, a variety of approaches like interviews, questionnaires, behavioral observation, and case studies (Robson, 1993) have been utilized in the process of mental health gathering information. However, within working and organization settings, occupational psychologists are often limited to observing natural variation, and therefore experimentation is difficult (Chmiel, 1998). As a consequence, in line with the accepted measuring techniques for personality, mental health studies have had preference for questionnaire-type approaches as a widely accepted measuring technique. However, as previously noted, issues regarding the use of self report inventories have been widely acknowledged.

## 3. SENSOR-BASED ESTIMATION METHODS

### 3.1 Affect Estimation Method

In a recent study, Egeuz Guevara and Umemuro (2010) utilized movement and infrared sensors to assess affect through movement descriptors such as walking speed, motion load, walking directness, and arms' movements. The subjects of the study performed daily activities in no-laboratory settings. In this method no devices were attached to subjects, and no emotion was induced. Behavior was assessed in terms of speed,

frequency and variability of human sensory data. This method was suited only for individual experimentation as subjects' identification was not available. Affect was estimated at a semi-real time basis where the margin to report the estimation was about a half to few hours.

### 3.2 Personality and Mental Health Estimation

In the method proposed by Eguez Guevara et al. (2011), sensory raw data were used to assess personality and mental health of subjects performing daily office routine. Acceleration and voice intensity data were captured by Business Microscope (Yano & Kuriyama, 2007) device which was worn by subjects as a name-tag. Sensory data corresponding to walking, talking, desk working, and idle behavior categories were analyzed in terms of variance, amplitude, or frequency and mapped to each listed behavior. It was then conducted a correlation analysis between the participants estimated behavior and their personality and mental health questionnaires scores.

### 3.3 Discussion

Both methods described in sections 3.1 and 3.2 surpass the invasive nature of conventional affect and personality and mental health measuring techniques which limits the normal activities and behavior of subjects being tested. Subjects of these methods contrast to conventional ones in that they need not to attend any laboratory setting, as the experimentation is done in their own living or working place. This prevents subjects from laboratories potential stress and distraction. Also, contrary to conventional techniques, the proposed in this section did not use any kind of made-up behavior or emotion. No induced-stimuli is beneficial as it has been argued that it is one reason for poor results in related studies for only weak, and context-lacking stimuli is generated.

Since the affect estimation method proposed in section 3.1 lacks of a subject identification system its experimentation in multi-subjects layouts is unfeasible. However, the time-to-output of both methods described in section 3.1 and 3.2 surpass the huge amount of time and resources needed by conventional estimation techniques. Furthermore, as these methodologies are set by continuously loading data, the estimated psychological states will always provide up-to-date information.

#### 3.3.1 Application

The use of sensory mechanisms in the measurement of psychological states brings broader opportunities to understand humans, and the relation between their behavior and psychological states.

For example, the availability of such information may help families' members to take proper actions in light of others' affect in order to enhance the individual capacity to provide affective support to other family members. For instance, a given family member *a*, who lives away from his/her family, might feel ease when he/she knows other family members appear with neutral or happy state. However, family member *a* may pay closer attention to his family condition, if other family members are found to be sad, or going through hard times. In this case, family member *a*'s awareness was achieved through the availability of the family members' psychological states and its adequate transmission.

From a different viewpoint, psychological states' information will also be informative to various research fields, for example to the study of the Affective Computing. Affective Computing, seeks to provide to any computer form the ability to recognize, express, regulate, and utilize emotions to respond to human emotion. Since personality, as any other human psychological state, is characteristic of human beings, the process of designing computers or robots which resemble to the human being should consider personality as an important aspect to characterize. The possibility of designing robots able to engender personality-based behavior opens discussions relating the advantages or disadvantages of the incorporation of these specific human characteristics.

## 4. CONCLUSION

The shift from explicit means of human input to more implicit forms of input enables more natural interactions with the physical environments, which provide sufficient input and information without demanding major burden to users. With the help of sensory technology, and without using physically demanding devices, or psychologically demanding monitoring techniques, it was suggested methodologies for the estimation of psychological states based on external signal like body movements and human behavior.

This paper opens a discussion for application possibilities that can make use of psychological states' information for the service and better understanding of human beings in non-critical environments. The accessibility to information of psychological states is of great interest to the development of services that ultimately are designed to serve people in more humane ways.

## REFERENCES

- Brave, S., & Nass, C. (2008). Emotion in human-computer interaction. In A. Sears & J. A. Jacko (Eds.), *The Human Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications* (2<sup>nd</sup> ed., pp. 77-92). New York: Taylor and Francis Group.
- Chmiel, N. (1998). Psychology in the workplace. In P. Scott & C. Spencer (Eds.), *Psychology: A contemporary introduction* (pp. 660-698). Oxford: Blackwell Publishers.
- Eguez Guevara, J., et al. (2011). Personality and mental health assessment: A sensor-based behavior analysis. *Proceedings of the Fourth International Conference on Advances in Computer-Human Interactions, Guadeloupe, France*, in print.
- Eguez Guevara, J. P., & Umemuro, H. (2010). Unobtrusive estimation of psychological states based on human movement observation. *eMinds: International Journal on Human-Computer Interaction*, 2(6), 39-54.
- Foner, L. N. (1996). A security architecture for multi-agent matchmaking. *Proceedings of the Second International Conference on Multiagent Systems, Japan*, 80-86.
- Fragopanagos, N., & Taylor, J. G. (2005). Emotion recognition in human-computer interaction. *Neural Networks*, 18(4), 389-405.
- Hutt, M. L. (1945). The use of projective methods of personality measurement in army medical installations. *Journal of Clinical Psychology*, 1, 134-140.
- John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of Personality* (pp. 102-138). New York: The Guilford Press.
- Meehl, P. E. (1945). The dynamics of "structured" personality tests. *Journal of Clinical Psychology*, 1, 296-303.
- Meyer, G. J., & Kurtz, J. E. (2006). Advancing personality assessment terminology: Time to retire "objective" and "projective" as personality test descriptors. *Journal of Personality Assessment*, 87(3), 223-225.
- Mischel, W. (1973). Toward a cognitive social learning reconceptualization of personality. *Psychological Review*, 80, 252-283.
- Picard, R. W. (1997). *Affective Computing*. Cambridge: The MIT Press.
- Robson, C. (1993). *Real World Research*. Oxford: Blackwell.
- Yano, K., & Kuriyama, H. (2007). Human x sensor: How sensor information will change human, organization, and society. *Hitachi Hyoron*, 89(7), 62-67.
- Wallbot H. G., & Scherer, K. R. (1989). Assessing emotion by questionnaire. In R. Plutchik and H. Kellerman (Eds.), *Emotion Theory, Research, and Experience: Vol. 4. The measurement of emotion* (pp. 55-82). New York: Academic Press.
- Wickens, C. D., & Hollands, J. G. (2000). *Engineering Psychology and Human Performance* (3<sup>th</sup> ed., pp. 17-68). Upper Saddle River, New Jersey: Prentice-Hall.
- Winter, D. G., & Barenbaum, N. B. (1999). History of modern personality theory and research. In L. A. Pervin & O. P. John (Eds.), *Handbook of Personality* (pp. 3-27). New York: The Guilford Press.
- Wilson, A. D. (2008). Sensor- and recognition-based input for interaction. In A. Sears & J. A. Jacko (Eds.), *The Human Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications* (2<sup>nd</sup> ed., pp. 177-199). New York: Taylor and Francis Group.

# PLAYING LIKE A GIRL: GENDERED GAMING AND GIRL GAMERS

Pirita Ihämäki

*Tampere University of Technology  
Unit of Human-Centred Technology  
Korkeakoulunkatu 6, 33101 Tampere, Finland*

## ABSTRACT

Massively multiplayer online games (MMOGs) are launched as the newest and fastest-growing genre of computer games; they could as much have been viewed as a return to the natural order of things. The advent of single-player genres as the central paradigm for games is an historical aberration of digital technology (Herz 1997). This paper examines digital games through gendered game content and game spaces. Study focuses on young women attraction the popular digital games and give proposes to design competitive game to reach female mass audiences. Underlying the position that there are fundamental differences between what boys and girls want from computer games is a discourse that posits essential differences in girls and boys interests and competencies. Female culture can be a significant extension into the realm of gaming, and contribute to the development of women's culture and the diversification of the gaming industry. Main conclusion is that gaming has to become a source of pleasure, mastery, and identity for girls.

## KEYWORDS

Gender, girl gamer, female player, girl's game

## 1. INTRODUCTION

On average, girls are less involved with digital games than are boys, and when they do play, they often prefer different games. Recent publications suggest that the amount of female game playing has increased, at least in the U.K. and the U.S. (Bryce & Rutter, 2002). According to current user data for the U.S. Market (ESA, 2005), 43% of all video game players are female. Online games form of digital game playing as well (44% of all online – players are female, ESA, 2005). The Sims success as a top selling digital game has been attributed to its attractiveness to female players (example Carr, 2005); this adds to the discussion of female-targeted game design (Graner-Ray, 2003; Gorriz & Mediana, 2000). Kafai & Giang (2009) examine girl players in Whyville.net, a virtual world, which have 68% of its players being female teens.

In sum, the research suggests that digital games have traditionally been a realm dominated by males, but that due to the emergence of female subcultures adopting contemporary digital games designed for males, and the advent of new games that successfully engage female players, the gender gap has started to narrow, at least in the U.S. Focusing on gender differences in involvement with computer games is primarily relevant for research on entertainment (e.g., Klimmt & Hartmann, 2006), as it seems plausible that gender-specific entertainment preferences exist. Gender is also an important factor in relation to use of digital games. Interestingly, massive multiplayer online role playing games (MMORPGs) seem to attract quite high percentages of female players (20-30 percent) and the research conducted by Taylor suggests that the social spheres provided by games like *EverQuest* can be more inviting and pleasurable for female players. (Hartmann, Klimmt, 2006.)

In this paper addresses the topic of gender and games and why young women are less attracted than young men to popular digital games. Second question is how to design competitive game if they intend to reach female mass audiences. In this article the research problems are of theoretical nature and my aspiration is to find solutions using concept analysis within the frameworks of digital culture and game culture. Concept analysis of gendered gaming and girl gamers is in the main focus of this article. Thus concept analysis can also be regarded as a kind of research approach. In the work of Ahonen and Kallio, defined by Näsi, concept



analysis is goal-oriented solving of concept problems, which produces new concepts and definitions. (Ahonen, Kallio 2002, 60.) Concept analysis utilises the method of reasoning, where both analysis and synthesis are applied to produce new concepts and frameworks. As a research approach conceptual analysis is purely theoretic, and is based on literature study. Based on literature, broader concepts are elaborated and deconstructed into smaller parts, to better understand their elements. (Neilimo, Näsi 1980, 31.) In this article gendered gaming and girl gamers raise up elaborated in game culture and try to find solution to reach female mass audiences in game culture. Observation of gaming and girl gamers in this article is based on other published articles. The approach in this work is descriptive. The purpose of descriptive study is to describe phenomenon, context or nature of an occurrence, its frequency, historical development or other qualities. (Uusitalo 1991, 62.) Materials used for this article are scientific articles and literature of the subject matter. I reflect upon the researchers' understandings of girl gamers and game fields in gender side. This introduction will map the range of girl gamer and female – target game design, offering a picture of girl gamer research. Study will deeply explore young women interest for digital games and try to fluid theoretical vision which suggests the design aspects to take account for designing new games on female players. This study will lead us further in the understanding of what computer games can be and what girls want from them. To conclude, this paper presents an organised description of development of aforementioned areas and future research.

## 2. RESEARCHER ON GENDER FOR GIRL PLAYERS'

"Gender" can be construed in different ways, depending on the researcher's disciplinary approach (see, e.g., for feminist and cultural-studies based approaches, Cassell & Jenkins, 1998); for psychological approach and overviews, (Bussey & Bandura, 1999). Bussey and Bandura (1999) theoretical perspective, gender roles and conceptions are the product of a broad network of social influences operating interdependently in a variety of societal subsystems. (p.676). Several authors have documented that most digital game heavily on stereotypes and archaic role models to portray female characters (Jansz & Martis, 2003). Female users could have low attraction and appreciation for such portrayals of female characters and their enjoyment and playing motivation could be reduced in consequence. While many of the woman playing online games are involved with more traditional games such as *Hearts of Bingo*, a growing number play MMOGs. Through the number of women do not outpace men, officials at three major MMOGs (*Asheron's Call*, *Ultima Online* and *EverQuest*) counted women as 20-30% of their subscriber base (Laber 2001). Female digital game players now comprise 40% of all players, and women over 18 make up more of the game playing population than do males under 17 (Top 10 Industry Facts, 2008). This genre offers a chance to revisit the question of women and gaming. (Taylor 2009, 94.)

However, it is important to recognize the multi-played nature of social life in spaces. The most basic understanding of online socialization frames the activity in terms of "chat" that you simply talk to people in the digital environment. Chatting, connecting with others, forming relationships and maintaining them are aspects of the interpersonal pleasure MMOG's afford and multiuser games have benefited by drawing in this component of digital games. (Taylor 2009, 94-95.) Talking about how women like the social component of games can flatten a fairly rich play landscape and trivialize the work involved in sustaining social life within a game. Identity exploration is also seen as a primary play goal for girls. For example in *EvertQuest* players has experimenting with creating selves. *EvertQuest* allows people to create up to eight distinct characters per server, means that is at least a potential to explore a range of persons. (Taylor 2009, 95.) Kerr (2005) study girls liked element of flexibility or freedom in games in terms of being able to explore the world in any order they liked and in relation to controlling the main character or creating their own character. Study shows that when girls had the time and were on their own they would choose a long, story driven single player game and when they had friends over or ended up playing games at a party it would be a competitive multiplayer game.

## 3. DESIGN COMPETITIVE GAME REACH FEMALE MASS AUDIENCE

While girl's new play games in increasing numbers, they spend less time playing games than boys and they play different kinds of games, factors that likely influence their motivation and opportunity to engage in technology-based practices associated with gaming. While some attention is being given by the game

development industry to including female game players into the games culture, when one went beyond the crude gender categories of games to explore what kinds of pleasures females got from playing games which they are interesting and emerged for them. Designing games that girls want to play is of course important, and designing games intended to increase girls' in development process. We need to find ways to provide girls with opportunities to engage digital games and participate in their communities. We need to take into account the social communities – be they virtual or face-to-face – that surround gaming and give certain IT practices values and significance. Multiplayer games like *World of Warcraft* are proving to be appealing to many women, and they offer many opportunities for technology-related learning. (Hayes, 2005) This research suggests that women may come to enjoy male-typed games, if given the time and support to master them. As Jenkins (2003) points out, games enjoyed by men often have features, such as avatar customization, that appeal to presumably “female” sensibilities. Games that combine elements associated with both stereotypically masculine and feminine pleasures and strengths may ultimately be the most stimulating and valuable games for entertainment. This leads to the issue of how we can encourage. Such communities, on-line as well as face-to-face, would provide opportunities for girls to develop multiple dimensions in games: not just technology skills, but also intellectual capabilities, conceptual understanding, engagement in social practices, (including collective knowledge building), and design knowledge. To accomplish these goals, communities need to be structured in ways to support multiple forms of legitimate participation, scaffolding not only in the use of tools but also in ways of thinking and problem-solving by more experienced peers or adults, and lots of talk, to develop fluency in communicating ideas and generally acquiring the identity of a individual. A particularly important aspect of these communities for girls is the opportunity to engage in “identity play,” finding ways to challenge their prior assumptions about the identities associated with digital expertise, to help them expand their conceptions of desirable “feminine” identities and practices, and to provide social recognition and legitimacy for these new identities.

#### 4. CONCLUSION

The role of active engagement in the game extends to the way one can interact with “worldness”. One of the notes in MMOG environments and many computer games offer is the way their construction of worlds lets users actually wonder a landscape and explore. Most women express a real enjoyment of engaging with the game as a world an environment. The issue of how virtual world experience “fitter back” is particularly striking, though, when woman players has reported that playing the game helped them become more confident or assertive. (Taylor 2009, 97.) That’s why games industry should consider how to implement competitive game structures if they intend to reach female mass audiences. As female player could create they avatars many ways but how about when players tries to imagine her place in the designers? The idea of the imagined user is key in shaping design and, by extension, the possibilities actual players encounter. From an industry perspective it is clear that the design and price of a console and the range of games available on it can act to deter or attract consumers. Female players wanted acknowledgement that players existed and flexibility in games so they were afforded more autonomy. As a practical consequence, the computer game industry stands to conquer new market segments if it would rigorously implement females’ content preferences in future products (Cassell, 2002; Gorritz & Medina, 2000) Overall, the findings contribute to an explanation of the substantial gender gap in computer game involvement. Result present that design the competitive game to reach female mass audiences, designers need to take account female needs. Female needs in digitals games are: rich social interaction tools, female wants to expression them identity examples create avatars, girls like flexibility or freedom in games. Girl players like also masculine games, like war game *World of Warcraft* if they get support of the gaming and they could become like hardcore gamers. This research conclusion is that girl player like also following other players more than male player and discuss for other player in game experiences. Discussing in game experiences is more important for girl players than male players, this elements is good to take in account for design the games. Researchers suggest for further theoretical discussion in entertainment research: current explanations of why playing digital games is fun need extensions to account for gender-specific models of pleasure. If competing, winning, and being a violent superhero do not appeal to women to the extent that they appeal to men, several mechanisms of enjoyment that have been proposed, such as pride in success. Most importantly, the pleasures of social interaction with game characters, and with other players as well, require more attention in theories about (interactive)

entertainment. In future it's interesting to ask what might be the impact on the general appeal of computer technologies to girls and women who are thinking about possible career path and participation in society (see Bussey & Bandura, 1999)?

## ACKNOWLEDGEMENT

I thank Marjo Mäenpää in Aalto University, who provided helpful comments on previous versions of this document.

## REFERENCES

- Ahonen, A. and Kallio, T. (2002) *Käsite- ja tekstitutkimuksen metodologia - perusteita, näkökulmia ja haasteita johtamis- ja organisaatiotutkimuksen kannalta (Methodolog in concept and text study, foundations, angles and challenges from the perspective of organisation study)*. Publications of Turku School of Economics. Series: Discussion and Reports, 5:2002. Grafia: Turku.
- Bryce, J. and Rutter, J. (2002) Killing like a Girl: Gendered Gaming and Girl Gamers' Visibility. In (Eds.) Mäyrä, F. *Computer Games and Digital Cultures, Tampere University Press, Tampere, 2002.*
- Bussey, K. and Bandura, A. (1999) Social cognitive theory of gender development and differentiation. *Psychological Review*, 106(4), p. 676-713.
- Carr, D. (2005) Context, gaming pleasures and gendered preferences. *Stimulation and Gaming*, 36(4), 464-482.
- Cassell, J. (2002) Genderizing HCI. In Jacko, J. & Sears, A. (Eds.) *The Handbook of Human-Computer Interaction*, p. 402-411.
- Cassell, J. and Jenkins, H. (Eds.). (1998) *From Barbie to Mortal Kombat: gender and computer games*. MIT Press, Cambridge, MA.
- Gorritz, C. M. and Medina, C. (2000) Engaging girls computers software games. *Communications of the ACM*, 43(1), p.42-49.
- Graner-Ray, S. (2003) *Gender-Inclusive Game Design: Expanding the Market*. Hingham, MA: Charles River Media.
- Hartmann, T. & Klimmt, C. (2006) Gender and Computer Games: Exploring Females' Dislikes. *Journal of Computer-Mediated Communication* 11 (2006).
- Hayes, E. (2005). Gendered identities at play. Paper presented at the Games, Learning, & Society Conference, Madison, WI, June 23-24.
- Herz, J.C. (1997) *Joystick Nation: How Videogames Ate Our Quarters, Won Our Hearts, and Rewired Our Minds*. New York: Little, Brown and Company.
- Jansz, J. and Martis, R. (2003) The representation of gender and ethnicity in digital interactive games. In Copier, M. & Raessens, J. (Eds.) *Level Up: Digital Games Research Conference*, p. 260-269. Utrecht: Utrecht University.
- Kafai, Y. B., et al. (2009) Transgressive Gender Play: Profiles and Portraits of Girl Players in a Tween Virtual World. *Breaking New Ground: Innovation in Games, Play, Practice and Theory*. Proceedings of DIGRA 2009.
- Kerr, A. (2003) *Girls/Women Just Want to Have Fun - A Study of Adult Female Players of Digital Games*. In: Level Up Conference Proceedings. University of Utrecht, Utrecht, p. 270-285.
- Klimmt, C. and Hartmann, T. (2006) Effectance, self-efficacy, and the motivation to play video games. In Vorderer, P. & Bryant, J. (Eds.) *Playing Video Games: Motives, Responses, and Consequences*, pp. 143-177.
- Laber, E. (2001) *Men are from Quake, Women are from Ultima*. The New York Times 11 January. <http://www.nytimes.com/>.
- Neilimo, K. and Näsi, J. (1980): *Nomoteettinen tutkimusote ja suomalainen yrityksen taloustiede: tutkimus positivismin soveltamisesta (Nomothetic Research Approach and Finnish Business Economics)*, Faculty of Economics and Administration, Series A2; Studies and Reports 12, Tampereen yliopisto.
- Taylor, T.L. (2009) *Play Between Worlds, Exploring Online Game Culture*, The MIT Press, Cambridge, Massachusetts, London, England.
- Top 10 Industry Facts. (2008). Washington, D.C.: Entertainment Software Association.
- Uusitalo, H. (1997) *Tiede, tutkimus ja tutkielma, Johdatus tutkielman maailmaan (Science, Research and Study an Interaction to Study, WSOY, Juva.*

# ENGINEERING ANONYMITY TO REDUCE AGGRESSION ONLINE

Michail Tsikerdekis

*Masaryk University, Faculty of Informatics, Brno*

## ABSTRACT

The effects of anonymity on aggression have been discussed by many social scientists in the past years. Anonymity is a factor that could lead to disinhibited behavior which could damage many online communities. This knowledge provides software engineers with a dilemma as to whether to use anonymity as an option for their users and suffer the increased aggression that their community might exhibit or exclude it all together. In order to cope with the increased aggression that might occur due to the presence of anonymity, most of the developers employ moderators to do policing work. This paper argues that there is a gap in the knowledge that we have about anonymity and aggression. Anonymity is a generic term and should be analyzed further into different states such as pseudonymity and complete anonymity. There is no knowledge covering the effects of these two different states of anonymity in the interactions of users in relevance to aggression. This paper shows that according to previous work there is enough evidence to support that there may be a difference. If this is true, software engineers might have another option for decreasing aggression by altering the design, instead of moderating their communities.

## KEYWORDS

Software, design, anonymity, aggression, pseudonym, interaction

## 1. INTRODUCTION

Anonymity as a factor that increases aggression online is well known by many observations that social scientists have made, but how much control software engineers have in order to affect and reduce anonymity's negative effects in the interactions of users is still a puzzling question.

Aggression as a destructive force has the ability to penetrate all the levels of communication within our lives and our digital life is no exception to that. On social networking media it is a powerful force which, given its ripple effects could propagate through networks and hold the potential to damage the reputation and the budget of corporations. In addition, damages are also accounted for by the victims that would have an unpleasant experience from their aggressor or aggressors. Cyberbullying can be damaging to individuals as an example and some of the incidents can be fatal for the victims which could even commit suicide, as this was the case of a female teenager that fell victim to a mother pretending to be a 16-year-old boy (Atkinson, 2008). So it becomes obvious that the danger is real for the victims but it is not hard to imagine why a company's reputation could be at stake if aggressive incidents persist within their community.

Many companies today that develop social networking media could consider increasing and improving supervision in order to reduce the violent incidents in their online communities, but healing the wound alone does not mean that the cause of the wound was eradicated as well. Moderation as an option might be a viable option as a policing method to reduce aggression but since it requires manpower it means that it comes with a cost. In addition moderation aims to deal mainly with the aftermaths of aggressive incidents rather than preventing them. In order to make an aggressive occurrence preventable, a shift in the way that software engineers are trained to think and work today is needed. Software engineers need to ask two questions about their software and their design interfaces. How does the software affect the interaction between the users and what can we do to change it?

Just as the same question needs to be asked for every project planning that has to do with social networking media, the same question could be imposed about anonymity and the effects that it has in the online interactions in relation to aggressive behavior. Ways of understanding these effects and altering the

software in order to achieve specific desired results through architectural design modifications, will be explored further in this paper.

## **2. THE CONCEPTS**

Before we move forward and see how these questions can be answered through simple experimentation we need to understand the key players in our equation, aggression and anonymity, as well as see how they can be analyzed and what we know so far about them when the two are combined.

### **2.1 Online Aggression**

Aggression online is something that can be seen in our everyday experience and interaction through computer-mediated communication. As in real life the same can be said about online aggression which in general is the delivery of an aversive stimulus from one person to another, with intent to harm and with an expectation of causing such harm, or perhaps when the other person is motivated to escape or avoid the stimulus (Russell G. Geen, 2001). Among human beings this can take a variety of forms such as physical, mental or verbal. Of course the physical one is excluded online due to the nature of online communication which renders it impossible, although there are examples with fights that occurred online and led to real world physical violence.

Another thing to consider is that since most of the online services for communication are for free, an individual can send an unlimited amount of emails, Facebook messages, tweets and other types of messages, and raise a significant havoc before administrators have the time to stop the individual. There are examples of leaked videos that were released online without someone's consent and the videos travelled all over the globe via online communication channels, even though administrators were trying to block them and delete their original videos. These cases should not only pose a privacy concern about the social media technologies and their impact to society but there should be a concern about why the individual chose to act in an aggressive manner and if technology provokes or affects this kind of behavior.

The occurrence of aggressive behavior is controlled by a number of situational factors and experiences. Examples of these factors can be the presence of violent objects such as a gun (Leonard Berkowitz and Anthony Le Page, 1967), or experiences like the military which provides the social context where servicemen learn aggression, violence, and murder (Castle, T.; Hensley, C., 2002). Among all the examples that show how situational factors can affect the aggressive response of an individual one stands out as probably the best example to show how circumstances can lead or even prevent aggression. That factor is frustration and the theory describing the effects of frustration on aggression is the frustration-aggression theory. The theory states that frustration can lead to anger, and that anger triggers a hostile action (Barker, R. Dembo, T., and Lewin, K., 1941). It is not hard to realize that online environments can be frustrating and aggression can be the product of this frustration. Even more important though, this theory clearly expresses the fact that aggression is heavily associated with situational factors and not just the psychosocial state and experiences of the individual. Hence, if these factors were to be understood, software engineers could develop the perfect environment which could provide the least amount of friction and better user experience for actually preventing aggression to a certain level.

### **2.2 Online Anonymity**

Anonymity refers to the state of an individual's personal identity, or personally identifiable information, being publicly unknown. Anonymity poses a couple of different issues concerning the informative effect, group pressure effect and enforcement effect with various pros and cons (Lee, G. B., 1996). There are numerous problems associated with anonymity such as hatred, harassment, violence, impersonation, libel, copyright infringement and more. But it also offers many advantages when it comes to reducing group pressure, allowing somebody to express himself without being penalized for it, and reducing potential discrimination due to someone's gender, age, ethnicity and other characteristics. Hence, discarding the ability to provide the users with the option of being anonymous on social networking media is not a straightforward decision for software engineers. As an example, research has shown that in votes and debates people voting

anonymously are more likely to change their vote and less likely to conform with the group's norm, behavior consistent with preventing groupthink type behavior which could lead to ineffective and even risky decisions (Ainsworth, S. et al., 2010). Providing the option of anonymity for a social networking brainstorming group is the right choice according to the above study.

On the other hand if the use of anonymity is being rendered as necessary by a software engineering team other ways of countering the negative effects of anonymity on the online community should be explored in order to counter them. The main effect of anonymity as a general term is that it is often associated with aggression. Even though an individual's identity might be traceable, simply the heightened feeling of anonymity appears to be enough to promote disinhibited behavior (Patricia Wallace, 1999). But that by itself is a general term especially for online situations were from a technical standpoint there is anonymity, unlinkability, linkability, undetectability, unobservability, pseudonymity and so on (Ptmann, A., Hansen, M., 2010). On the other hand from a user's standpoint anonymity could probably be divided into three different levels, anonymity being absent when the user uses his or her real name, pseudonymity when a nickname is being used, and complete anonymity when there is no trace evidence to find out who the individual is. This paper tries to pinpoint the differences between the levels when it comes to these three states of anonymity in relation to aggression. The issues that arise when it comes to deciding which one to use and why we cannot reach a satisfying conclusion with the information that we have so far.

### **2.2.1 Factors for Disinhibited Behavior and Complete Anonymity**

According to the online disinhibition effect, a term that was invented to describe the way that people online feel less restrained and express themselves more openly, toxic disinhibition (the negative side of disinhibition) is affected by a number of factors such as dissociative anonymity, invisibility, asynchronicity, solipsistic introjection, dissociative imagination, and minimization of authority (John Suler, 2004). The factors that are interesting for this paper and are more associated with anonymity are the dissociative anonymity but also the dissociative imagination.

This provides us with an important answer for the state of a completely anonymous user. The first factor dissociative anonymity, which describes the state of anonymity that disassociates the individuals from their real names, can justify disinhibited behavior under the state of the user being completely anonymous. Dissociative anonymity is so strong that an individual might even convince him or herself that he or she has no responsibility for the online actions.

### **2.2.2 Pseudonyms & Disconnecting with the Online Self**

When it comes to the use of pseudonyms there seem to be more than two factors at play. Aside from the dissociative anonymity there is also the second factor of dissociative imagination which describes the idea of one depicting that their online alter ego exists beyond the realm of reality and therefore it could act in a different manner than one's real self. It is easy to understand how dissociative imagination is associated more with the use of characters with nicknames/pseudonyms and not with complete anonymity since there needs to be a character creation process in order to act as a factor.

Dissociative anonymity and dissociative imagination might have a significant effect on the disinhibition effect between individuals that use pseudonymity and individuals that use complete anonymity. According to the above remarks an individual could grow an attachment to his or her pseudonym self and become more aggressive just because the same rules that apply in the real world do not apply in the online world. On the other hand a user with complete anonymity is lacking the factor of dissociative imagination and therefore his actions might be less aggressive.

### **2.2.3 Pseudonyms & Connecting with the Online Self**

On the opposite side of the seesaw research on IRC nicknames and impression formation seems to suggest that nicknames are an inherent part of their Net-identity, and even of their "real-life" identity (Bechar-Israeli, H., 1995). People find the need to describe their traits, characteristics, appearances with their nicknames and try to find the optimum way to do so. In a way they try to encapsulate part of their personality into a nickname creating an online extension of their real selves and not a completely different and independent alter ego.

In fact the research that was mentioned above shows that people grow a long term attachment with their nicknames and usually prefer to keep the same nickname and identity which for the most part is connected

with the real self which they wish to exhibit. In addition most of the people when selecting a nickname do not base their decisions in collective values but rather values related to the self. Finally many people in their real lives have the tendency to call others by their online nicknames, another sign that nicknames have a strong connection to our real lives.

#### **2.2.4 Disconnecting versus Connecting with the Online Self**

There are two conclusions that can be extracted from the above remarks about online anonymity and aggression. The first, is that complete anonymity is a substantial factor that could lead to disinhibited behavior and aggression. The second conclusion is not as straight forward as the first one and brings us to a contradiction. When it comes to pseudonymity by using nicknames/pseudonyms there are two powers at play. The first is that people try to include part of their personality into the online persona that they are creating while the second, dissociative imagination, tries to drive the creating process beyond the realm of reality. The question is which one of these two powers wins over and therefore how pseudonymity differs from complete anonymity in relation to aggression.

The above remarks could hinder an important revelation about pseudonymity and complete anonymity when it comes to aggression. There is a chance that aggression levels could differ between users that use pseudonyms and completely anonymous users. The difference being that users with pseudonyms might be more polite than the completely anonymous ones, because there is a strong connection with the self and therefore disinhibited behavior is reduced to a certain level, or users that use pseudonyms can be more aggressive than users that are completely anonymous. There is also a chance that there may be no difference between the two.

From a design perspective the benefits of knowing these differences for software engineers can be extremely important because by simple changes in the design, i.e. shifting from completely anonymous users to users with pseudonyms or the vice versa, one can have a significant effect on the aggression level that a community exhibits through its users. In addition if aggression can be reduced by these alterations, moderation costs for social networking media could also decrease.

### **3. CONCLUSION**

The concepts discussed in this paper show the potential benefits of understanding the effects of the design on the user interaction especially for the case of anonymity and aggression. We know that anonymity is a contributor for aggression but anonymity as a general term is not enough to describe all the different states of anonymity and the factors that affect each one of these cases. This paper has also shown that even though there seems to be a difference in the effects of pseudonymity and complete anonymity because of the different factors that affect them, the difference cannot be determined from a theoretical standpoint since pseudonymity could have various effects on an individual.

There is a gap in our current knowledge of anonymity and aggression which needs to be explored. Pseudonymity could play a key role for reducing aggression online or increasing it. Finding out what is the directionality of pseudonymity in relation to aggression is important for software engineers.

That was the reason for the decision being made to conduct an experiment which could qualitatively and quantitatively measure the effects of different anonymity states in the aggression levels of individuals. The experiment being run is currently active, and tests on how individuals respond in various scenarios based on different controversial topics in three different anonymity states. The states are the pseudonymous, completely anonymous and the control state where anonymity is absent, meaning that the users are using their real names. Therefore any statistically significant differences can be observed between the different anonymity states. In addition, for each topic individuals were asked about how strong their opinion is about the different topics in order to see if users that have a stronger opinion about topics are more susceptible to the effects of different anonymity states.

The outcome of this research will provide us with results that could have real implications in the way that social networking software is being designed today. The aim is to decrease violent behavior such as cyberbullying that could be attributed to the effects of different anonymity states in online communities where debates take place. If a different anonymity state can be used to decrease the amount of violent incidents the benefits are immense for the social networking media. In this paper, the economic benefits for a social

networking media developer have been made obvious, but there is also the potential for making safer communities for the users. That benefit might be more important than any economic advantage that this research may yield. Since the goal of the social networking software is to serve the interactions of users, understanding how these interactions relate to the interface and the architectural design is essential for making the future of online communication safer, more pleasant and more effective.

## ACKNOWLEDGEMENT

First and foremost, I would like to thank my supervisor, Jiri Zlatuska for the valuable guidance and advice which made this research possible. In addition, I would like to thank the Faculty of Informatics, Masaryk University of Brno for helping fund the research and for the precious help that I received from the staff of the Faculty.

## REFERENCES

- Ainsworth, S. et al., 2010. Anonymity in classroom voting and debating. *Learning and Instruction*, Vol. 21, No. 3, pp. 365-378. doi:10.1016/j.learninstruc.2010.05.001
- Barker, R. Dembo, T., and Lewin, K. (1941). Frustration and aggression: An experiment with young children. *University of Iowa Studies in child Welfare*, Vol. 18, pp 1-314.
- Bechar-Israeli, H., 1995. FROM ‹Bonehead› TO ‹cLoNehEAd› : NICKNAMES, PLAY, AND IDENTITY ON INTERNET RELAY CHAT. *Journal of Computer-Mediated Communication*, Vol 1, pp 0. doi: 10.1111/j.1083-6101.1995.tb00325.x
- Berkowitz, L., LePage A., 1967. Weapons as aggression-eliciting stimuli. *Journal of Personality and Social Psychology*, Vol. 7, No. 2, pp 202-207.
- Castle, T.; Hensley, C. 2002. Serial Killers with Military Experience: Applying Learning Theory to Serial Murder. *International Journal of Offender Therapy and Comparative Criminology*, Vol 46, No. 4, pp 453-65.
- John Suler, 2004. The Online Disinhibition Effect. *CyberPsychology & Behavior*, Vol.7, No. 3, pp. 321-326.
- Lee, G. B., 1996. Addressing Anonymous Messages in Cyberspace. *Journal of Computer-Mediated Communication*, Vol 2, No. 1, pp 0. doi:10.1111/j.1083-6101.1996.tb00183.x
- Patricia Wallace, 1999. *The psychology of the internet*. Cambridge University Press, Cambridge, UK.
- Ptzmann, A., Hansen, M., 2010. Anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management (A consolidated proposal for terminology). [http://dud.inf.tu-dresden.de/Anon\\_Terminology.shtml](http://dud.inf.tu-dresden.de/Anon_Terminology.shtml). Version 0.34.
- Ronald Atkinson, 2008. Societal effects of cyberbullying: the dark side of building bridges with technology. *South Atlantic Philoshopy of Education Society Yearbook*.
- Russell G. Geen, 2001. *Human aggression*. Open University Press, Buckingham – Philadelphia, USA.





# Posters



# STUDY ON DICHROMATISM'S COLOR APPEARANCE TO LED LIGHTS

Mamoru Takamatsu, Yoshio Nakashima<sup>1</sup>, Tetsuya Fujita<sup>1</sup> and Kazuhiro Sawa<sup>2</sup>

<sup>1</sup>*University of Toyama*

*3190 Gofuku Toyama-city, 930-8555 Japan*

<sup>2</sup>*Nihon Color Technology Research*

*1-14-6-721 Tomita-chou, Takatsuki-city, Osaka 569-0814 Japan*

## ABSTRACT

In recent years, with the utilization of blue Light Emitting Diode, the Light Emitting Diode element is beginning to be actively used in various scenes, and Light Emitting Diode is used even for traffic light now. The Light Emitting Diode traffic signal light has little power consumption as compared with the traffic signal light using the conventional electric bulb, and spreads to progress increasingly from now on. Because the discernment of the yellow of a signal light and red is difficult for sense-of-color unusual person, in the present signal light system, yellow signal light color is made somewhat brighter than red signal light. The aim of the present study is to collect the fundamental data for the display of the traffic signal light, which is easy to be recognized also by sense-of-color unusual persons. The results revealed that the optimal brightness of the yellow light from which a sense-of-color unusual person can distinguish the difference between yellow and red, i.e., "visual barrier-free domain", is the area where yellow light is 2.8 times brightness of the red light.

## KEYWORDS

Color appearance, Dichromatism, LED lights

## 1. INTRODUCTION

In accordance with the fact that the blue LED came into practical use in these years, wide-range color display became also possible in the LED display unit by combining it with traditional red and green LED's. Currently, LED devices have actively been utilized in various locales. For instance, LED type road information boards have rapidly been diffusing in the field of land transportation. Recently further, the LED type traffic signal light has also been developed and its number of installation is rapidly growing. That has excelled property in visibility, particularly in preventive effect for pseudo-lighting due to rising or setting sun, which has been longtime subject of bulb-type traffic signal, so its expectation for preventing effect against accident is extremely high in the viewpoint of safety. There are also many merits up to prevention of earth warming caused by energy saving in addition to cost reduction for maintenance and checkup due to its long lifetime. As mentioned like this, the LED-type traffic signal has many merits and is expected to spread more and more from now, whereas it has various matters in question.

It has currently been said that number of the chromatic disabled in Japan is ca. 3 millions, rated to ca. 5% in male and 0.02% in female<sup>1</sup>). Viewing state by the chromatic disabled is due to aberration in any of visual substances corresponding to 'red', 'green' and 'blue' on the retina, and symptom gets largely different depending on the visual substance. By the way, case of abnormality on visual substance for 'red', that for 'green' and that for 'blue' are called as the 1st, 2nd and 3rd disorder in chromatic vision, respectively (e.g., Pitt 1935; Wright 1946; Wright 1952). The 2nd disorder occupies the largest fraction among these, and the 1st and 2nd disorders are generically called red-green chromatic disorder because their symptoms are mutually alike, and the most of the chromatic disabled belongs to this red-green disorder.

To this kind of chromatic disabled people, when information on the premise of color discrimination is given, reading the information there from is accompanied by an extreme difficulty. In the traffic signal, in particular, it gives a great influence on safety, too, because the yellow and red signals are hardly distinguished. Accordingly, the present signal light has coped with it by making color of the yellow signal a

little brighter than the red light, but there are extremely few examples of study that verified if this system is truly effective. Now then, the present study set its purpose to collect basic data for presenting the traffic signal easy to distinguish even by the chromatic disabled.

## 2. EXPERIMENTAL PROCEDURE

The present study purposes, particularly for the traffic signal that has played an important role to maintain the traffic safety, to get the optimum luminance for yellow and red signal lights for distinguishing them easily even by the dyschromatopic. Prepared first are 'yellow stimulus light' corresponding to the yellow signal and 'red stimulus light' to the red signal by using an image preparation software. Luminance value of the red reference light is  $14\text{cd/m}^2$ , so in the yellow stimulus light, we prepare 11 kinds of luminance values in 0.2-time step in the range between 1.0 to 3.0 times compared with the red reference. The reference light and posing stimulus light prepared are converted into viewing state of the intense 2nd chromatic disabled by means of chromatic simulation software (Brettel et al. 1997). By simultaneously posing the yellow stimulus light after the conversion and the red reference light, person to be tested does judge if difference from the red reference can be distinguished, when varying luminance value of the yellow arbitrarily, with the 3-step assessment standard. With this sort of procedure to simulate the viewing state for the chromatic disabled, we enabled an evaluation experiment by using the chromatic normal as the subject.

Experimental procedure is shown as follows: The subject is first subject to dark adaptation for 10 minutes in a dark room. Then, yellow and red posing lights (which have been converted into the viewing state by the intense 2nd chromatic disabled) are posed on the LED posing apparatus for 2 sec as shown in Fig. 1. Task of the subject is to observe 2 posed lights with both eyes and to answer if difference of the yellow stimulus light (left) with respect to the red reference light (right) can be distinguished with the 3-step assessment standard.

Namely, the subject does assess with 3 steps of '0' for indistinguishable case, '1' for the case decidable neither, and '2' for distinguishable case. Next, blank (dark screen) is posed for 4 seconds. By following with similarity, the image that changes luminance of the yellow stimulus light at random and the blank are alternately posed, and the subject answers whether he could distinguish or not.

Besides, the experiment was performed in a dark room in order to exclude stray light as much as possible. Background color in posing the stimulus was set black, and size of the circular posing stimulus light on the posing unit was set at  $2^\circ$  in viewing angle. The subjects were 10 people of the chromatic normal and measurement was conducted 5 times per person. Enter the text here.

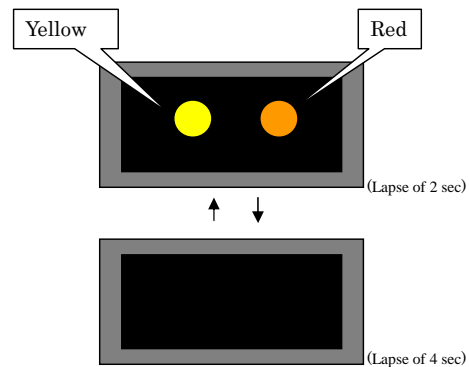


Figure 1. Picture plane of LED display unit

## 3. CONCLUSION

An example of the result is shown in Fig. 2. It shows a graph with luminance value of the yellow stimulus light (scale factors in parentheses are those of luminance of the yellow stimulus with respect to that of the red) as abscissa and selectivity of the assessment '2' (distinguishable) as ordinate. Standard deviation of the selectivity at each luminance value has been shown by overlapping as the error bar in the graph.

It has been shown from the result that selectivity is extremely low and value of error bar tends to stay small in the range where luminance value of the yellow stimulus light is 1.0 to 1.6 times as much as that of the red. This fact has indicated that these 2 colors easily distinguishable with hue, even if having nearly equal luminance, for the chromatic normal are extremely indistinguishable for the chromatic disabled. Whereas in the range where luminance value of the yellow stimulus light is 1.8 to 2.2 times as much as that of the red, value of error bar was large and showed passable fluctuation though selectivity got a little increased. Further

in the range up to 2.6 times, value of the error bar results in staying large though the selectivity has rapidly increased. Namely, it is found under conditions of these ranges with difference in luminance that they are as vague as 'they can neither be distinguished or not' being unknown. In the range where luminance value of the yellow is from 2.8 to 3.0 times as much as the red, selectivity becomes very high and further, value of the error bar is also small, in other words, scattering has become not so much visible. This is found to be the range where difference from the red is easily distinguished also by the chromatic disabled.

On the other hand, Fig. 3 shows relationship between variation in luminance of the yellow and the 3-step evaluation of selectivity. Similarly to Fig. 3, it has taken variation in luminance value ( $\text{cd}/\text{m}^2$ ) of the yellow on abscissa while the selectivity (%) of assessment '0', '1' and '2' on ordinate. By the way, standard deviation has been omitted here to make easier to see the result.

In the range where luminance value of the yellow stimulus light is 1.0 to 2.0 times as much as the red reference light, assessments '0' (indistinguishable) and '1' (decidable neither) have occupied the most of selectivity, and this region is found also from this Fig. as the range that 'yellow and red colors are indistinguishable or hardly distinguished'. On the other hand, fraction of assessment '2' (distinguishable) starts rising when luminance value of the yellow gets beyond 2.0 times, and between both sides of the cross point of graphs of assessment '1' (indistinguishable) and '2' (distinguishable), the relationship of the two graphs reverses. It is considered therefore that use of the range below 2.4 times is better to be avoided as much as possible.

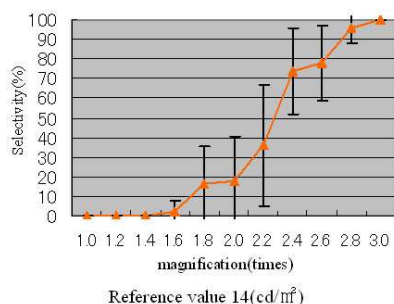


Figure 2. Transition of selectivity for assessment '2' with luminance.

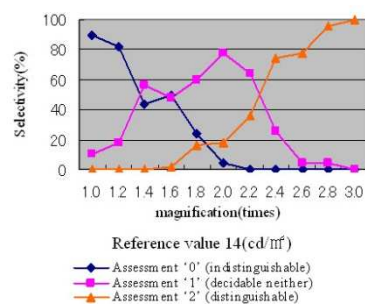


Figure 3. Electivity for each item of assessment.

The present study purposes to derive the optimum luminance of display for the traffic signal that is easily perceivable by the chromatic disabled. From the result of approximated curve obtained, we got the knowledge that it is the most appropriate to set luminance of the yellow traffic signal above 2.8 times with respect to that of the red one for the display easily observed even by the chromatic disabled.

## REFERENCES

- Brettel, H. Vienot, F. John Mollon, D., 1997. Computerized simulation of color appearance of dichromats. *J. Opt. Soc. Am.*, 14(10), 2647.
- Pitt, F. H. G., 1935. Characteristics of dichromatic vision, *Med. Res. Council. Spec. Repser. Nr.*, 200.
- Wright, W.D., 1946. Researches on Normal and Defective Color Vision, *Kimpton, London*.
- Wright, W.D., 1952. The characteristics of the tritanoptia, *J. Opt. Soc. Am.*, 42, 509-521.

# COLOR SENSITIVITY TO LED LIGHTS IN THE ELDERLY

Yoshio Nakashima<sup>1</sup>, Mamoru Takamatsu<sup>1</sup>, Hiroki Fujita<sup>1</sup>, Kazuo Sendai<sup>2</sup> and Santarou Nakajima<sup>3</sup>

<sup>1</sup>*University of Toyama, 3190 Gofuku Toyama-city, 930-8555 Japan*

<sup>2</sup>*Panasonic Electric Works Consulting & Training Co., Ltd., 1048 Oojikadoma, Kadoma-city, Osaka 571-8686 Japan*

<sup>3</sup>*Seiwa Electric MFG. Co., Ltd., 36 Terada-Shinike Johyou-city, Kyoto 610-0192 Japan*

## ABSTRACT

Aging is rapidly progressing in Japan, and the number of patients of senile cataract is steadily increasing. This fact has not only become a medical subject but also a large social problem. Because manner that the cataract people get a view is different from that of the normal who are not so, some consideration is considered necessary. Fraction of cataract people has reportedly been near 60% for ages of the 50's and beyond 90% for the 80's or higher, so it may be said that countermeasure against this sort of visual impairment is also one of extremely important subjects in order to raise life quality of the elderly. In these years, by the way, we have become visible of many LED-type articles for the train departure information within stations, the designation display of buses and the like. LED has been collecting eyes to merits of long lifetime and low power consumption, and its number of installation and range of application are considered to increase further. However, the present state of reported examples regarding the manner of viewing in chromatic vision of the elder, particularly for the LED lighting display board, is extremely few. In this experiment, the authors attempted study and speculation with respect to the manner of viewing colors on the multicolored LED lighting display board by means of goggle that fictionally reproduce visual environment of the cataract. It became evident from the result that perception of color has a large difference from yellowish green to aqua between visual environments of the cataract and normal, and further we qualitatively revealed its range and difference, too.

## KEYWORDS

Color sensitivity, elderly, LED lights

## 1. INTRODUCTION

In the rapidly growing birthrate-falling and aging society, at present, number of the senile cataract is steadily increasing year after year. In most cases, cataract appears from aging variation of crystalline lens in the eyeball, and is one of senile symptoms that can come off to everybody same as white hair and wrinkled skin. In cataract, external image gets dim and hardly visible because light becomes harder to enter oculus due to opacity of the crystalline lens. Namely, manner of viewing comes slightly different from healthy youth.

Proportion of the cataract disabled in terms of age has been regarded near 60% for the 50's and over 90% for the 80's, so it may be said that measures for this sort of the visually disabled caused by cataract are one of very important subjects in Japanese society where aging is rapidly progressing. The aging society stands for rise in number of people with some physical handicap in other words, and accordingly it is important to grow barrier-freeing in various aspects, so we will also be quotable "visual barrier-free" as one of them.

In these years, by the way, we have become visible of many LED-type articles for the train departure information within stations, the designation display of buses and the like. LED has been collecting eyes to merits of long lifetime and low power consumption, and its number of installation and range of application are considered to increase further. However, the present state of reported examples regarding the manner of viewing in chromatic vision of the elderly (e.g., Van den Berg, T. J., 1995; Ikeda et al., 2003), particularly for the LED lighting display board (e.g., Takamatsu et al. 2001; Tsuru et al., 2002), are extremely few.

In the present study, we purposed to contemplate and take consideration regarding how colors are seen in the cataract visual environment by employing multicolored LED lighting bulletin board and cataract-environment reproducing goggle (The Government of Ontario, Canada) from the perspective that collects basic data for information board display eye-friendly for people of every generation from the elderly with a relatively high population of the cataract to the juvenile that are not so.

## 2. EXPERIMENTAL PROCEDURE

In this experiment, we used a multicolored LED lighting display board as the stimulus posing device. With ceiling lamps in the laboratory room and auxiliary fluorescent lamps, illuminance on surface of the lighting display was set constant to 300 lx, which corresponds to that of departure guidance board actually installed at the station.

Experimental procedure is shown below. At first, a circular stimulus light with viewing angle of  $2^\circ$  (ca. 14 cm dia.) is posed on a lighting display board against the black background for 2 sec. The subject observes this test stimulus light at a spot 4-m apart (Fig. 1) under the condition of wearing a goggle reproducing cataract visual environment, which corresponds to a level of considerable intensity. Task of the subject is to freely answer the color name perceived from the test stimulus light posed, namely, with “free naming method”, and by this we intended also to reveal on elemental sensory colors (categorical colors) (e.g., Ratcliff 1976; Boynton 1980; Boynton and Olson 1987; Boynton and Paul 1989) in parallel. In consequence, 4-sec blank is set before posing the following stimulus light for excluding, as much as possible, influence of color conformance with the test stimulus light that has been posed. Subject’s answer with free naming is conducted during this 4-sec blank in practice. Similarly hereinafter, the test light and blank are alternately posed, and the subject answers color name perceived for each stimulus light whenever posed. This is also tried with similar procedure in the case using no goggle. By the way, sequence of giving the test stimulus light was set so as to be at random for every subject and trial.

Stimulus lights used for the test are 78 colors, measured so as to make color difference even on sides and inside the triangle having its apex at the 3 points of R, G and B with the highest chroma expressible on a multicolored LED lighting display board. Luminance of LED display is 105 cd/cm<sup>2</sup> for R ( $u' = 0.517$ ,  $v' = 0.522$ ), 124cd/cm<sup>2</sup> for G ( $u' = 0.07$ ,  $v' = 0.567$ ), and 48.5cd/cm<sup>2</sup> for B ( $u' = 0.164$ ,  $v' = 0.144$ ), under the condition of monochromatic presentation. On the other hand, luminance of the background was 25cd/m<sup>2</sup>. Defining test stimulus light posing of these 78 colors for each subject as one session, we tried 10 sessions in total, comprising each 5 sessions in the state with the goggle and that without, respectively. By the way, the subjects are 10 people with normal chromatic vision.

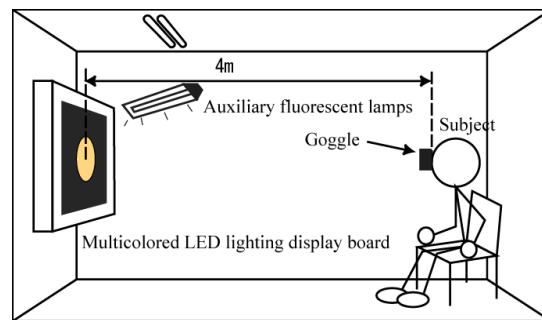


Figure 1. Experimental situation

## 3. RESULTS AND CONCLUSIONS

Shown in Figure 2 are collective plots of points answered with the color name at appearance probability above 50% for each of totally 78 test stimulus lights as the range of respective colors. Results in the state with and without goggle that reproduces cataract visual environment are (a) and (b), respectively, while (c) is a plot of only overlapped stimulus points answered with the same color name from (a) and (b).

From results in (a) and (b), number of obtained colors, namely categorical colors, were 10, comprising red, orange, yellow, yellowish green, green, aqua, purple, pink and white. When comparing results of the both, however, we observed change of some colors in each color range. With respect to yellowish green, green and aqua, in particular, the range perceived as the color name showed trend to get narrow rapidly under cataract visual environment (a) in comparison with the case being not so (b). In 2 colors of yellow and white, moreover, the range has largely shifted to the left and it was suggested that the juvenile and senile possibly perceive as separate color names, even when posing the identical chromatic point, in colors in these ranges along with colors mentioned above.

Further from the result of (c) showing the range where (a) and (b) overlap, there is no overlap between the range perceived yellow by the senile and that by the juvenile for the yellow color. Accordingly, special attention is considered necessary in order to use this hue. In addition, ranges are extremely narrow with respect to aqua, yellowish green and white, too, and when using these 3 colors, it is considered hopeful not to



largely deviate from chroma points of  $(u', v') = (0.115, 0.374)$  for aqua,  $(u', v') = (0.122, 0.562)$  for yellowish green and  $(u', v') = (0.178, 0.444)$  for white.

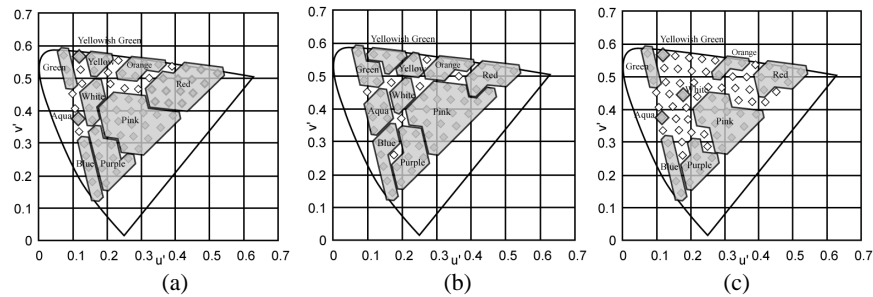


Figure 2. Collective plots of points answered

In the present study, we attempted to collect basic data on how color of the multicolored LED lighting display board under visual environment for cataract by means of goggle to reproduce it.

From the result, regarding the color perception of the elderly with many cataracts and that of the juvenile being not so, the authors qualitatively revealed it including variation of the range. With respect to 4 categorical colors of yellow, yellowish green, aqua and white, sufficient attention was suggested necessary for use of them.

## REFERENCES

- Boynton, R. M., 1980. Categorical colour perception colour rendering of light source. *Proc. CIE (Venice)*, 66-69.
- Boynton, R. M. and Olson, C. X., 1987. Locating basic colors in the OSA space. *Color Res. Applic.*, 12, 94-105.
- Boynton, R. M. and Paul, K. F., 1989. Categorical colour perception under low-pressure sodium lighting with small amounts of added incandescent illumination. *Lighting Res. Technol.*, 21, 23-27.
- Ikeda, M. Kusumi, A. Obama, T. and Shinoda, H., 2003. Color appearance of color charts observed with a cataract experiencing goggle. *Journal of the Color Science Association of Japan*, 27(2), 113-124.
- Ratcliff, F., 1976. On the psychophysical bases of universal colour names. *Proc. Amer. Phil. Soc.*, 120, 311-330.
- Takamatsu, M. Nakashima, Y. Nakajima, S. Katoh, Z. Yoshihara, S. and Iizuka, M., 2001. Optimum number of display colors in a short time of presentation -in case of LED traffic information board-. *Japanese Journal of Visual Science*, 22(1), 7-10.

# AN INTRODUCTORY STUDY ON PLACEMENT METHOD OF BACKGROUND MUSIC FOR SHOPPING WEBSITE

Chien-Jung Lai<sup>1</sup>, Wen-Ming Wu<sup>1</sup>, Ya-Ling Wu<sup>1</sup> and You-Shyang Chen<sup>2</sup>

<sup>1</sup>*National Chin-Yi University of Technology*

*35, Lane 215, Section 1, Chung-Shan Road, Taiping 411, Taichung, Taiwan*

<sup>2</sup>*Hwa Hsia Institute of Technology*

*111, Gong Jhuan Rd., Chung Ho, Taipei, Taiwan*

## ABSTRACT

The main purposes of this study are to design placement methods of background music for shopping website and examine the effect of these methods on browsers' emotional and cognitive response. An online questionnaire survey was conducted first to understand the website browsers' behavior and attitude toward background music. Then a laboratory experiment will be conducted. Results of the online survey indicated that 65.5% of browsers noticed if there is background music on websites. 52.6% of browsers could accept the existence of background music on website. However, 61% of browsers considered that it is not necessary to place background music on websites. Two placement methods of background music (i.e., placement point and fade-in period) for shopping website have been proposed and will be undertaken in a laboratory experiment. The results could help researchers to know the theoretical explanations of how and why these placement methods influence browsers, and help marketer to adopted valid placement methods of background music for an online store.

## KEYWORDS

Placement method, background music, shopping website.

## 1. INTRODUCTION

Online retailing has attracted a great deal of attention in recent years due to the rapid developing of the Internet. Some researchers have been already begun to call for more systematic research by using established retailing and consumer behavior theories. As a natural departure from the stimuli present in a traditional retail store, the online retail environment can only be manipulated by visual and auditory cues. In the past, research of online store focused on the design of website structure and interface from visual stimulus, few carried on the discussion from the auditory stimulus. Recently many websites place background music to attract browsers' attention. Some researches start to study the effect of background music on consumer response (Wu et al., 2008). These researches assumed the existence of background music. However, this premise is not necessarily tenable. Browser masters the broadcast of background music of online website. He or she may stop broadcast momentarily. Then the effects of background music have no way to begin with. To discuss the effect of background music of online store, it is necessary to allow browsers to accept the existence of background music first.

## 2. RESEARCH ON STORE ATMOSPHERICS

The impact of atmospherics on the nature and outcomes of shopping has been examined by researchers for some time. To explain the influence of atmospheres on consumer, atmospheric research has focused heavily on the Mehrabian-Russel affect Model (Mehrabian and Russell, 1974). Empirical work in the area has examined specific atmospheric cues and their effects on shopper response in brick-and-mortar environment. As in brick-and-mortar environments, atmospheric cues have been posited to influence consumers on the web (Eroglu et al., 2001). Dailey (2004) defined web atmospherics as "the conscious designing of web

environments to create positive effects in user in order to increase favorable consumer responses (e.g., site revisiting, browsing, etc.)". Eroglu et al. (2001) classified web atmospheric cues into high task-relevant cues (i.e., descriptions of the merchandise, the price, etc.) and low task-relevant cues (i.e., the colors, typetypes and fonts, music and sounds, etc.). These cues form the atmosphere of a web site. However, research on web atmospheric thus far is limited in its theoretical explanation of why web atmospherics influence consumers. Dailey (2004) indicated that web atmospheric researchers should begin to focus on specific web atmospheric cues and theoretical explanations of how and why these cues may influence consumers.

There have been a number of studies that investigate the effect of background music on the physical environment (Herrington and Capella, 1994; Oakes, 2000). Bitner (1992) argued that music is a critical ambient condition of the servicescape and that music influences people's emotions and physiological feelings, and mood. Various structure characteristics of music, such as time (rhythm, tempo, and phrasing), pitch (melody, keys, mode, and harmony), and texture (timbre, orchestration, and volume), influence consumer response and behavior (Bruner II, 1990). However, few studies have reported the effect on website (Wu et al., 2008). The present study will address this issue by focusing specifically on design of music cues for shopping website.

### **3. RESEARCH METHODOLOGY**

This study conducted an online questionnaire survey first to understand the website browsers' behavior and attitude toward background music. Then two placement methods of background music are proposed and will be examined in a laboratory experiment.

#### **3.1 Questionnaire Survey**

A total of 449 valid online questionnaires were collected in December 2010. Results indicated that 65.5% of browsers noticed if there is background music on websites. 52.6% of browsers could accept the existence of background music on website. However, 61% of browsers considered that it is not necessary to place background music for websites. Browsers considered that the background music is suitable for the content websites (41.4%) and community websites (30.5%). Pop music (57.7%) and non-lyrics music (69.3%) are preferred as background music. The analysis of results also showed that background music has no significant effect on browsers' cognitive response, such as attention, browsing period and the intention of browsing.

#### **3.2 Laboratory Experiment**

This study proposes two placement methods of background music for shopping website. The first method considers the placement point of background music during the browsing. Three entry points from the start of browsing (i.e., 2 min., 4 min., and 6 min) are considered. The second method is the fade-in period of background music. The fade-in period of background music refers to the time interval of music volume increasing gradually from zero to a certain volume. 2 min., 4 min., and 6 min. of fad-in period are considered in the experiment. Fig. 1 shows the conceptual framework of the two placement methods. The maximum volume is set to be 60db. A 3(placement point) X 3(fade-in period) between-subject experiment will be performed to explore the effects of these methods on browsers' response. Both browsing without music and browsing with constant music volume will be treated as control groups. An online book store is designed and serves as the context of this experiment. A classical concerto with fast tempo is used for background music. Participants are asked to browse the website and then complete a purchase task. Variable amount of time to browse is allowed as long as the participants want. Mehrabian and Russell's 12-item semantic differential scale (1974) will be employed as the emotional response measure. Participants' recognition memory of online store contents, perceived time and attention (eye fixation time and frequency, and eye track orbit, EEG record) are collected for the analysis of browsers' cognitive response measures. It is important to know these cognitive responses for the design of online store.

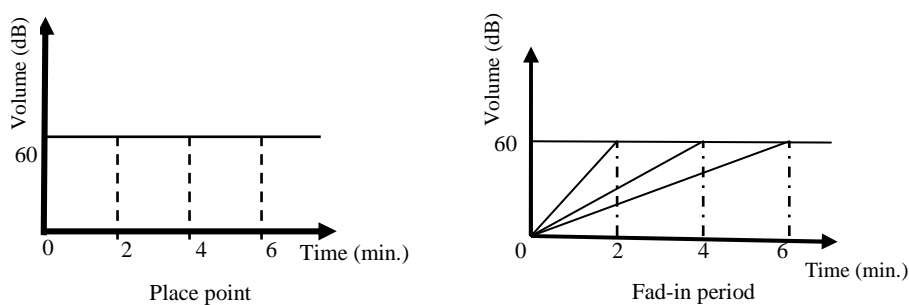


Figure 1. The conceptual framework of the two placement methods for background music

#### 4. CONCLUSION

This study is to design the background music placement methods and examine effect of these methods on browsers' emotional and cognitive response. Two placement methods of background music, i.e., placement point and fade-in period for shopping website have been proposed. The effects of these placement methods will be examined in a laboratory experiment. The results could help researchers to know the theoretical explanations of how and why these methods influence browsers. Another implication of this study is to help marketer to manipulated suitable placement modes of background music for a web store.

#### ACKNOWLEDGEMENT

This research was funded by the National Science Council of Republic of China under contract no. NSC 99-2221-E-167-020-MY2.

#### REFERENCES

- Bitner, M. J., 1992. Servicescape: The impact of physical surroundings on consumers and employees, *Journal of Marketing*, Vol. 56, pp. 57-71.
- Bruner II, G.C., 1990. Music, mood, and marketing. *Journal of Marketing*, Vol. 54, No. 4, pp. 94-104.
- Dailey, L., 2004. Navigational web atmospherics explaining the influence of restrictive navigation cues, *Journal of Business Research*, Vol. 57, pp. 795-803.
- Eroglu, S. A., Machleit, K. A., and Davis, L. M., 2001. Atmospheric qualities of online retailing: A conceptual model and implications. *Journal of Business Research*, Vol. 54, No. 2, pp. 177-184.
- Herrington, J.D. and Capella, L.M., 1994. Practical application of music in service settings, *Journal of Services Marketing*, 8, 3, pp. 50-65.
- Mehrabian, A. and Russel, J. A., 1974. *An Approach to Environmental Psychology*. MIT Press, Cambridge, MA.
- Oakes, S., 2000. The influence of the musicscape within service environments. *Journal of Services Marketing*, Vol. 14, pp. 539-550.
- Wu, C. S., Cheng, F. F. and Yen, D. C., 2008. The atmospheric factors of online storefront environment design: An empirical experiment in Taiwan. *Information & Management*, Vol. 45, pp. 493-498.

# PETRI NET BASED HCI SPECIFICATION AND CONTROL FOR VIRTUAL MAINTENANCE TRAINING

Wen-Yang Feng<sup>1</sup>, Chin-Jung Chao<sup>1</sup>, Feng-Yi Tseng<sup>1</sup> and Tien-Lung Sun<sup>2</sup>  
<sup>1</sup>*Chung-Shan Institute of Science and Technology, Long-Tan, Tao-Yuan 325, Taiwan*  
<sup>2</sup>*Department of Industrial Engineering and Management, Yuan-Ze University*  
*135 Yuan-Tung Rd., Chung-Li, Tao-Yuan 320, Taiwan*

## ABSTRACT

Due to its rigorous mathematical foundation and intuitive graphical representation, Petri Net (PN) has been used by researchers to model the human-computer interactions (HCI) in a virtual environment (VE). In manufacturing domain, PN is also employed by researchers to model HCI in virtual product support training systems. Previous researches, however, combined the low-level interactions, i.e., object manipulations, with the high-level control logics like task coordination and decision-making strategies. The resulted PN is difficult to be constructed and maintained, especially for complicated interaction scenarios. To address this problem, we divide the HCI control program of a virtual training system into three levels and develop PN models at each level. We also link the PN models to the VE to control the executable behavior scripts in the VE. In the future, the proposed PN model will be used to build a virtual maintenance training system for a military vehicle and the performance evaluation will also be conducted.

## KEYWORDS

Virtual reality, HCI modeling, Petri net, maintenance training.

## 1. INTRODUCTION

Product support training like operation or maintenance training is important for products with high cost, long life time and complicated structures. As 'hands-on' training with the physical product is usually expensive and dangerous, virtual reality (VR) technology has been proposed as a promising tool for low-cost and safe training. In general, a VR-based product support training system consists of three functional modules: graphic model, domain knowledge module and pedagogical module. The graphic model defines the geometry and appearance of the virtual objects; the domain knowledge contains instructions and constraints about the training task; and the pedagogical module monitors the trainees' behavior and evaluates their performance. The graphic model is built using 3D modeling and animation tools or 3D CAD systems. After the graphic model is built, it is imported to a VR or game authoring tool in which the interactions between the virtual objects and the trainee are programmed according to the rules and constraints defined in the domain knowledge model and pedagogical module.

## 2. PETRI NET BASED HCI CONTROL

The property that Petri Net (PN) has both rigorous mathematical foundation and intuitive graphical representation makes it an attracting tool for modeling and controlling the human-computer interactions (HCI) in VE (Willans and Harrison 2001, Ying 2006, Rieder et al. 2010). The ability to describe the dynamic behavior of a complicated system in a graphic representation makes PN a suitable tool to define the interactions in a virtual maintenance training environment. The graphic notation also makes the PN model easier for domain experts and VR engineers to define, modify, and maintain the interaction scenarios.

Many researchers have employed PN for HCI control in VR-based product support training. Lin et al. (2002) employed PN to represent the machining operation procedures for operation training. The PN model developed for operation training, however, can not be directly applied to maintenance training because the

tasks performed are different. In operation training, the tasks performed are simple operations like button pushing, cover lifting, etc.. In maintenance training, the tasks performed are more complicated disassembly or assembly operations. Several researchers (Kashiwa *et al.* 1995, Ishii *et al.* 1998) employed PN to control HCI in maintenance training. But the PNs developed combined the low-level HCI control activity, i.e., object manipulation control, with the high-level HCI control activities like task coordination and planning. The resulted PN is difficult to be constructed if the training scenario modeled is large and complicated.

### **3. HIERARCHICAL PN FOR HCI CONTROL IN MAINTENANCE TRAINING**

Raposo *et al.* (2001) proposed a three-level hierarchy for collaborative virtual environment design. The highest level, workflow level, describes the global sketch of the environment's behaviour. It specifies the actors and the tasks assigned to the actors. Each interdependent task at the workflow level is expanded into a middle level coordination task. The lowest level, called execution level, specifies the actual execution of tasks.

The 3-level hierarchy proposed by Raposo *et al.* (2001) is adopted in this paper to classify HCI control activities in a VR-based maintenance training system. Below we describe the PN models developed at each level.

#### **3.1 Execution PN**

The PN model built at the execution level controls the object manipulations in the VE, e.g., button pushing; cover opening, part disassembly/assembly, and so on. In a desktop VR environment, these operations are simulated using mouse clicking, mouse drag-drop, and so on. In a more advanced VR setting where data gloves and 3D trackers are available, object manipulations could be simulated using more natural way like hand grabbing and releasing. Sun (2009) has described an execution PN to control the unexpected situations in object manipulation.

#### **3.2 Coordination PN**

The PN developed at the coordination level handles resource and temporal management in a maintenance training task. For example, if a disassembly task requires the trainee to use one hand to hold a handle object and the other hand to grab a tool to loosen a screw, then the HCI program has to monitor the trainee's actions to make sure that the 'handle holding' task is finished before the 'screw loosening' task starts. Also the HCI program has to check whether the trainee performs the 'tool grabbing' task before he/she performs the 'screw loosening' task. As another example, if a training task needs to be done by the cooperation of two or more people, then the HCI program has to monitor the actions performed by multi-users to coordinate the resources and the temporal relationships.

In this work, the PN proposed in (van der Aalst *et al.* 1994) is adopted to build the coordination PN for one-hand disassembly (or assembly) task. The PN proposed in (Raposo *et al.*, 2001) is adopted to build the coordination PN for two-hand disassembly task.

#### **3.3 Workflow PN**

The PN model developed at the workflow level dynamically determines the sequence of training tasks. For a constrained training, the training tasks performed have a fixed sequence, e.g., following the standard operation procedure defined in the technical manual. The corresponding PN model thus has a simple linear structure. For a more flexible training, e.g., the exploratory type of training (Bluemel *et al.* 2003), the HCI control program must allow the trainee to explore different ways to conduct a task. In this case, the PN model has a more complicated structure like the And-Or graph or the disassembly constraint graph as proposed in (Li *et al.* 2003) if a more flexible training is required.

#### 4. LINKING THE PN MODEL WITH THE VR MODEL

The workflow PN is not linked to the VE as it only controls the execution of the next two levels of PN. The coordination and execution PNs are linked to the VE. At the writing of this paper, we are still investigating the details to link the coordination PN with the VE. The link between the execution PN and the VE has been studied by (Sun 2009). The places define the different disassembly status. They send control commands to modify the behavior scripts, e.g., increasing or decreasing the speed for a part to follow the mouse. The transitions store rules that define the firing conditions, e.g., the disassembly status of a part based on its transition or rotation data, whether a part is grabbed by the mouse, and so on.

#### ACKNOWLEDGEMENT

This research is partially supported by National Science Council, Taiwan, R.O.C. under contract NSC 99-2221-E-155-021-MY3. The author appreciates the financial support of NSC, R.O.C.

#### REFERENCES

- van der Aalst, W. M. P., van Hee, K. M., and Houben, G. J. 1994. Modelling and analysing workflow using a Petri-net based approach. *Proceeding of the 2nd Workshop on Computer-Supported Cooperative Work, Petri nets and related formalisms*. pp. 31-50.
- Bluemel, E., Hintze, A., Schulz, T., Schumann, M. and Stuering, S., 2003. Virtual environments for the training of maintenance and service tasks. *Proceedings of Winter Simulation Conference*, New Orleans, LA, Vol 2, pp. 2001-2007.
- Ishii, H., Tezuka, T., and Yoshikawa, H., 1998. A study on design support system for constructing machine-maintenance training environment based on virtual reality technology. *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, Vol. 3, pp. 2635-2640.
- Kashiwa, K. I., Mitani, T., Tezuka, T., and Yoshikawa, H., 1995. Development of machine-maintenance training system in virtual environment. *Proceedings of the IEEE International Workshop on Robot and Human Communication*, pp. 295-300.
- Li, J. R., Khoo, L. P. and Tor, S. B., 2003. Desktop virtual reality for maintenance training: An object oriented prototype system (V-REALISM). *Computers in Industry*, Vol. 52, No. 2, pp 109-125.
- Lin, F., Ye, L., Duffy, V. G. and Su, C. J., 2002. Developing virtual environments for industrial training. *Information Science*, Vol. 140, No. 1-2, pp 153-170.
- Raposo, A. B., Cruz, A. J. A., Adriano, C. M., and Magalhaes, L. P., 2001. Coordination components for collaborative virtual environments, *Computers and Graphics*, Vol. 25, pp. 1025-1039.
- Rieder, R., Raposo, A. B. and Pinho, M. S., 2010, A methodology to specify three-dimensional interaction using Petri Nets, *Journal of Visual Languages and Computing*, Vol. 21, pp. 136–156.
- Sun, T. L., 2009. Petri Net-based VR model interactive behavior specification and control for maintaining training. *International Journal of Computer Integrated Manufacturing*, Vol. 22, No. 2, pp. 129–137.
- Willans, J. S. and Harrison, M. D. 2001. A toolset supported approach for designing and testing virtual environment interaction techniques, *International Journal of Human-Computer Studies*, Vol. 55, pp. 145-165.
- Ying, J., 2006. An Approach to Petri Net Based Formal Modeling of User Interactions from X3D Content. *Proceedings of the 11th International Conference on 3D Web Technology*, pp.153-157.

# VISUAL AND MANUAL PERFORMANCE OF HUMAN COMPUTER INTERACTION IN VIRTUAL ENVIRONMENT

Chiuhsiang Joe Lin<sup>1</sup>, Hung-Jen Chen<sup>1</sup> and Tien-Lung Sun<sup>2</sup>

<sup>1</sup>*Department of Industrial and System Engineering, Chung-Yuan Christian University  
200, Chung Pei Rd., Chung-Li, Taiwan 32023, R.O.C.*

<sup>2</sup>*Department of Industrial Engineering and Management, Yuan-Ze University  
135, Yuan-Tung Rd., Chung-Li, Taiwan 32003, R.O.C.*

## ABSTRACT

Recent advancement in virtual reality (VR) technology has made possible the human user high sense of immersion in interacting with the VR environment. It is well known that even under real world the human is subject to perception errors. In VR environment, it is very likely that the perception of distance would affect how well the human user move and control the body in order to work in VR environment. The proposed study aims to explore the factors that affect the visual perception of distance in VR environment and thus understand how these visual factors further affect the manual interaction performance. The present study will be divided into two experiments. The purpose of the first experiment is to investigate the distance perception factors, especially the depth perception which has been proposed as the main problem in many VR systems. Other factors include the visual angle in the visual field and how far the observer is from the object. The refresh rate may also play important roles in distance perception. The measures will include the precision of the distance evaluation and the subjective discomfort in various VR parameters. The second experiment is to investigate the manual control performance in VR environment and its relationship with the distance perception. The Fitts' law task will be used as the experimental task and the participant's manual movement speed and accuracy will be measured under various VR parameters. In particular, to control the experimental size under management, the movement task will be carried out under the worst and best conditions in some of the factors investigated by the first experiment. The findings will provide a basis for designing a better VR environment for highly accurate manual control in that environment, that is, a better understanding in human-VR interaction.

## KEYWORDS

Virtual Reality, Depth Perception, Movement Performance.

## 1. INTRODUCTION

Recent advancement in virtual reality (VR) technology has made possible the human user high sense of immersion in interacting with the VR environment. It is well known that even under real world the human is subject to perception errors. In VR environment, it is very likely that the perception of distance would affect how well the human user move and control the body in order to work in VR environment.

Human depth perception in general is based on different sources of information. The impact of cue dominance depends on the observer's situation and on the distance between observer and object (Cutting & Vishton, 1995). When looking at distances up to 5 meters, the visual depth sensitivity of accommodation, convergence, and binocular disparity drops, whereas the latter has the highest impact (Nagata, 1993).

Several studies on distance perception in virtual environments using head-mounted displays (HMDs) have found significant underestimation of egocentric distance, the distance from an observer to a target - in some cases up to 50% or more, relative to distance perception in the real world (Messing & Durgin, 2005; Witmer & Sadowski, 1998). Previous studies have ruled out a significant role for a wide variety of physical factors, such as the limited field of view available in a HMD vs. in the real world (Knapp & Loomis, 2004), and although Willemsen, et al. (2004) have found evidence that issues related to the ergonomics of head mounted display systems may account for some of the apparent compression observed, an explanation for the larger portion of the effect remains lacking. In the most significant recent finding, Thompson, et al. (2004)



have demonstrated that the “graphical quality” of the virtual environment also cannot account for the observed effect.

The underestimation of distance was also reported in Armbruster, et al. (2008) study in which the projection screen was used as a VE display method. The accuracy of verbal estimation decreased as object distance increased. Subsequently, Naceri, et al. (2009) conducted an experiment where subjects had to compare relative depths of two virtual objects. The experiment was performed with both HMDs and immersive wide screen displays using the same visual cues to estimate depth. The experiment results showed significant differences between the two devices and confirm the distance misestimation phenomenon for HMDs.

In addition to the physical factors motioned above, visual angles seem to be a determinant factor for relative distance perception in real world. Errors in perceived relative distance diminished as visual angle increased. (Matsushima, et al., 2005).

Although the relative distance perception in real world has been studied, few experiments have directly investigated visual angle in VEs. In addition, limited research has documented refresh rate as a crucial factor to depth perception in VEs. Systemically exploring the factors that affect the visual perception of distance in VEs will not only supplement the finding of these early studies but also enable effective improvement of future VR technique.

For this reason, the present study aims to systemically explore the factors that affect the visual perception of distance in VR environment and thus understand how these visual factors further affect the manual interaction performance. The present study will be divided into two experiments. The purpose of the first experiment is to investigate the distance perception factors, especially the depth perception which has been proposed as the main problem in many VR systems. Other factors include the visual angle in the visual field and how far the observer is from the object. The refresh rate may also play important roles in distance perception. The measures will include the precision of the distance evaluation and the subjective discomfort in various VR parameters. The second experiment is to investigate the manual control performance in VR environment and its relationship with the distance perception. The Fitts’ law task will be used as the experimental task and the participant’s manual movement speed and accuracy will be measured under various VR parameters. In particular, to control the experimental size under management, the movement task will be carried out under the worst and best conditions in some of the factors investigated by the first experiment. The findings will provide a basis for designing a better VR environment for highly accurate manual control in that environment, that is, a better understanding in human-VR interaction.

## 2. METHODS

The experiment is a within subjects experiment. The factors are two experience environments (virtual environment presented with ViewSonic PJD6251 3D projector and real environment), five horizontal visual angles (60°, 75°, 90°, 105°, and 120°, Fig. 1), three vertical visual angles (15°, 30°, and 45°, Fig. 2), four distances between observer and targets (50, 100, 150, and 200 cm, Fig. 3), and three refresh rates (30Hz, 60Hz, and 90Hz). The verbal estimation values (in cm) and Simulator Sickness Questionnaire (SSQ) (Kennedy, et al., 1993) scores are the dependent variables. Estimations, which lay 10% over or under the scaled distance, will be categorized as hits; values below this range as underestimations; and values above this range as overestimations (Armbruster, et al., 2008). In addition, SSQ ratings responded to each condition will be collected at the end of the experiment.

At the start of each condition, subjects will be allowed practice trials to familiarize tasks. Following this, subjects will carry out all 360 experimental conditions in a random order to minimize the effect of learning. Then the estimation trials start. First, two spherical targets will be presented simultaneously apart from each other for 1 cm to 10 cm randomly in each experimental condition (as shown in Fig. 4), and participants have to verbally estimate the distances, in centimeters, between two targets. The experimenter will note the estimation results. No cues or time restrictions will be provided. After the experimental task, subjects will rate the SSQ assessments.

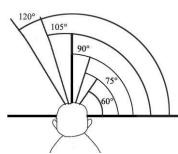


Figure 1. Horizontal visual angles

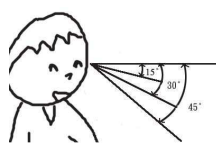


Figure 2. Vertical visual angles

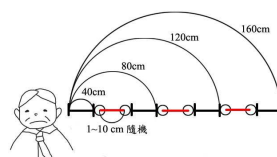


Figure 3. Distances between observer and targets

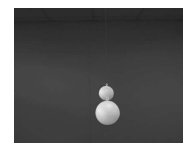


Figure 4. Relative distance estimation tasks

### 3. EXPECTED RESULTS

The present study aims to systemically explore the factors that affect the visual perception of distance in VR environment and thus understand how these visual factors further affect the manual interaction performance during human and virtual environment interaction. Based on the results from our experiments, visual estimation errors and manual control errors in the virtual environment can thus be modeled. The predicted data can be used to predicted the behavior of human users and improve the design of human computer interaction in the virtual reality scenarios.

### ACKNOWLEDGEMENT

This paper was supported by National Science Council of Taiwan under contract No. NSC 98-2221-E-033-020-MY2.

### REFERENCES

- Armbruster, C., Wolter, M., Kuhlen, T., Spijkers, W., & Fimm, B., 2008. Depth perception in virtual reality: distance estimations in peri- and extrapersonal space. *Cyberpsychology Behavior*, Vol. 11, No. 1, pp 9-15.
- Cutting, J. E., & Vishton, P. M., 1995. *Perceiving layout: the integration, relative potency, and contextual use of different information about depth* (Vol. 5). San Diego: Academic Press.
- Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G., 1993. Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology*, Vol. 3, No. 3, pp 203-220.
- Knapp, J. M., & Loomis, J. M., 2004. Limited Field of View of Head-Mounted Displays Is Not the Cause of Distance Underestimation in Virtual Environments. *Presence*, Vol. 13, No. 5, pp 572-577.
- Matsushima, E. H., de Oliveira, A. P., Ribeiro-Filho, N. P., & Da Silva, J. A., 2005. Visual angle as determinant factor for relative distance perception. *Psicológica*, Vol. 26, pp 97-104.
- Messing, R., & Durgin, F. H., 2005. Distance Perception and the Visual Horizon in Head-Mounted Displays. *ACM Transactions on Applied Perception* Vol. 2, No. 3, pp 234-250.
- Naciri, A., Chellali, R., Dionnet, F., & Toma, S., 2009. Depth Perception within Virtual Environments: A Comparative Study Between Wide Screen Stereoscopic Displays and Head Mounted Devices. *Paper presented at the Future Computing, Service Computation, Cognitive, Adaptive, Content, Patterns*, 2009. COMPUTATIONWORLD '09. Computation World.
- Nagata, S., 1993. *How to reinforce perception of depth in single two-dimensional pictures*. In S. Ellis (Ed.), *Pictorial communication in virtual and real environments* (pp. 527-545). London: Taylor & Francis.
- Thompson, W. B., Willemsen, P., Gooch, A. A., Creem-Regehr, S. H., Loomis, J. M., & Beall, A. C. (2004). Does the Quality of the Computer Graphics Matter when Judging Distances in Visually Immersive Environments? *Presence*, Vol. 13, No. 5, pp 560-571.
- Willemsen, P., Colton, M. B., Creem-Regehr, S. H., & Thompson, W. B., 2004. The effects of head-mounted display mechanics on distance judgments in virtual environments. *Paper presented at the Proceedings of the 1st Symposium on Applied perception in graphics and visualization*.
- Witmer, B. G., & Sadowski, W. J., 1998. Nonvisually Guided Locomotion to a Previously Viewed Target in Real and Virtual Environments. *Human Factors*, Vol. 40, pp 478-488.

# HUMAN PULSE WAVE MEASUREMENT BY MEMS-ECM SENSOR

Shusaku Nomura<sup>1</sup>, Masako Hasegawa-Ohira<sup>1</sup>, Yasushi Hanasaka<sup>1</sup>, Tadashi Ishiguro<sup>2</sup>  
and Hiroshi Ogawa<sup>3</sup>

<sup>1</sup>*Nagaoka University of Technology, 1603-1 Kamitomioka, Nagaoka, Niigata, 940-2188 Japan*

<sup>2</sup>*TAIYO YUDEN Co., Ltd., 6-16-20 Ueno, Taito, Tokyo, 110-0005 Japan*

<sup>3</sup>*Bifröstec Inc., 2-38 Kandajinboh, Chiyoda, Tokyo, 101-0051 Japan*

## ABSTRACT

A novel schema for human pulse wave measurement using millimeter-sized silicon microphone is proposed in this study. With regard to the microphone, an Electret Condenser Microphone (ECM) fabricated by Micro Electro Mechanical System (MEMS) technology was employed. We investigated the frequency characteristic of the MEMS-ECM to see if such small microphone originally designed to match the auditory frequency band had enough sensitivity in the lower frequency band at around 1.0Hz; this infrasonic sound is a dominant frequency of the human pulse wave generated pressure difference. As a result of sound sweeping assessment, it was found that the MEMS-ECM possessed a high sensitivity in the lower frequency band at around 1.0Hz even though the frequency characteristic was in the form of the differentiation circuit in such lower frequency band. Based on this result, we proposed possible variations of the pulse wave measurement introducing this MEMS-ECM.

## KEYWORDS

Pulse wave, MEMS, ECM, silicon microphone.

## 1. INTRODUCTION

A micro Electret Condenser Microphone (ECM) fabricated by Micro Electro Mechanical System (MEMS) technology was employed as a novel apparatus for human pulse wave measurement. Since ECM frequency response characteristic, i.e. sensitivity, logically maintains a constant level at lower than the resonance frequency (stiffness control), the slightest pressure difference at around 1.0Hz generated by human pulse wave is expected to measure by ECM. However, with regard to MEMS-ECM as a single device, the difficulty arises in a possible sensitivity reduction due to a high input-impedance amplifier embedded in MEMS-ECM. Therefore as a foremost task, it should begin with the verification of the frequency response characteristic at around pulse wave, which is dominant at 1.0Hz.

## 2. MEMS-ECM FREQUENCY RESPONSE CHARACTERISTIC AND PULSE WAVE MEASUREMENT

A consumer MEMS-ECM, SPM0408HD5 provided by Knowles Acoustics, LLC., USA/Japan, is used in this study. The frequency response characteristic of this MEMS-ECM in the band from 0.125 Hz to 100 Hz was obtained by FFT analyzer (CF-7200, ONO SOKKI Co., Ltd., Japan); by which sine wave sweeping signal was generated and given to an ordinary dynamic speaker in which MEMS-ECM was attached onto its diaphragm. As a result of the experiment, MEMS-ECM has found to have -20dB/dec of sensitivity reduction in that frequency band. Thus it is assumed to be equivalent with a differentiation circuit in which a single pole at around 100 Hz exists. Therefore the pulse wave taken by this MEMS-ECM should be in the form of differentiation, or velocity.

As shown in Fig.1, introducing compensation circuit, human pulse wave (Fig. 1(a)) was successfully obtained by integrating MEMS-ECM raw signal (Fig. 1(b): pulse velocity), and also that in acceleration form was successfully obtained by differentiating the raw signal (Fig. 1(c)).

### 3. POSSIBLE APPLICATIONS AND CONCLUSION

The strong directional characteristics of MEMS-ECM make a remarkable difference in possible applications of this device. Fig. 2 shows an example of the artery tracing by MEMS-ECM. Since it is a micro device with strong directional characteristics, it possesses high spatial resolution. In fact, the plotted pattern on the surface of forearm where pulse wave signal was detected successfully traces the radial artery, which is one of two main arteries embedded in the forearm. This unique profile is resulted by the high spatial resolution of MEMS-ECM. Moreover since MEMS-ECM is produced in the semiconductor production process, each sensor is superior in quality. Therefore it is easy and preferable than other sensors to assemble multiple measurement systems.

In this study, it is successfully demonstrated that MEMS-ECM can be a novel and useful apparatus for human pulse wave measurement. In the field of clinical medicine and health science, hemodynamic characteristics of the body provide us with useful and important information for maintaining our daily health.

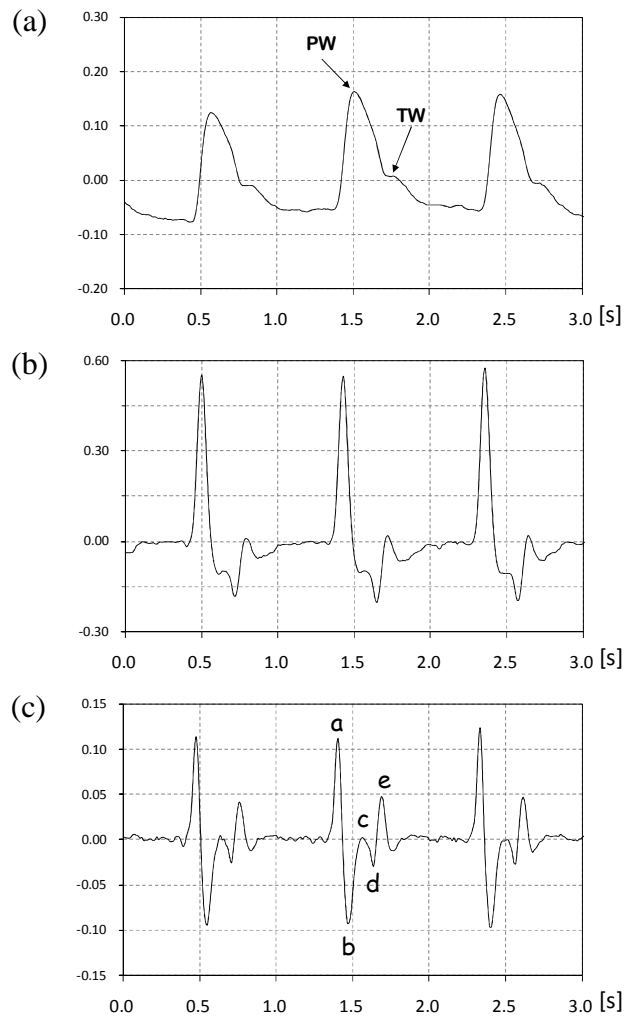


Figure 1. (a) Pulse wave, (b) pulse wave (velocity), and (c) pulse wave (acceleration) obtained by MEMS-ECM.

Moreover such hemodynamic characteristics has been investigated in the field of ergonomics, human interface, and other related engineering since it is known to have close relationship with the emotional and affective state of human. Far more exploitation of the use of this micro device for the measurement of human cardiovascular system toward health care is expected.

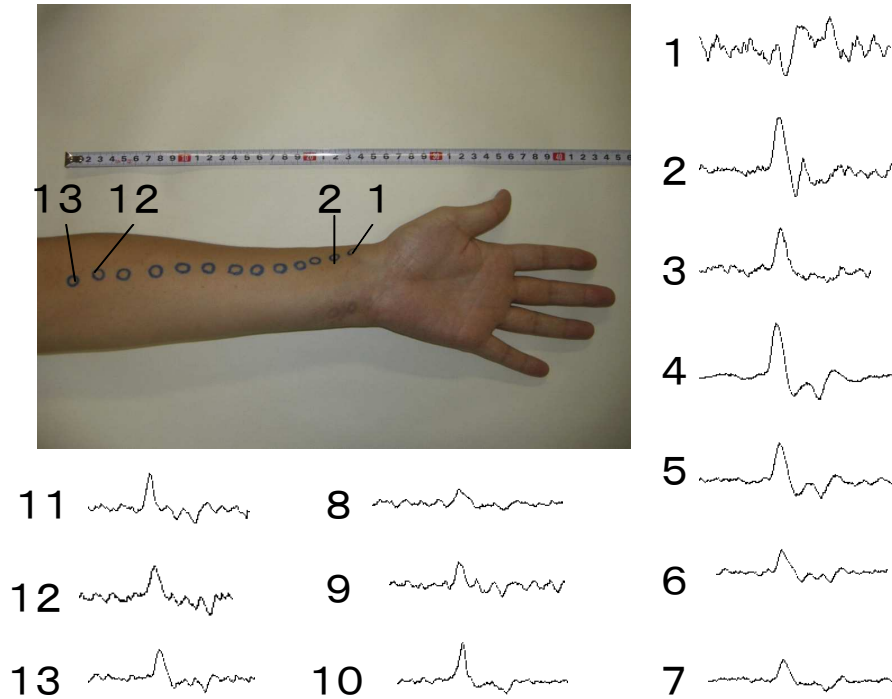


Figure 2. Artery pulse tracing by MEMS-ECM.

## REFERENCES

- W. Chen, K. Kobayashi, S. Ichikawa, Y. Takeuchi, and T. Togawa: "Continuous estimation of systolic blood pressure using the pulse arrival time and intermittent calibration", *Med Biol Eng Comput.*, Vol.38, No.5, pp.569-574 (2000)
- Healey JA and Picard RW: "Detecting Stress During Real-World Driving Tasks Using Physiological Sensors", *IEEE Trans. on Intelligent transportation systems*, Vol.6, No.2, pp.156-166 (2005).
- Williams RB: "Patterns of reactivity and stress; Handbook of stress, reactivity, and cardiovascular disease (K.A. Matthews, S.M. Weiss, T. Detre, T.M. Demobroski, B. Falkner, S.B. Manuck, R.B. Williams Eds.)", John Wiley & Sons (1986).

# EVALUATION OF WCAG 2.0 SUCCESS CRITERIA ON LUMINANCE CONTRAST RATIO BASED ON AGE-RELATED SPECTRAL EFFICACY

Tetstuya Fujita<sup>1</sup>, Yoshio Nakashima<sup>2</sup> and Mamoru Takamatsu<sup>2</sup>

<sup>1</sup>*Faculty of Art and Design, University of Toyama, Japan*

<sup>2</sup>*Graduate School of Science and Engineering, University of Toyama, Japan*

## ABSTRACT

To evaluate WCAG2.0 success criteria on luminance contrast ratio, we executed the precise measurement of spectral radiance for web-safe colors, and calculated the age-related relative luminance and contrast based on the age-related spectral luminous efficacy, which is defined in Japanese Industrial Standard. As a result, reduction of luminance by aging is observed, especially among people in their 70s, at the bluish region. Also, by calculating the age-related luminance contrast ratio against black background, we have found out that 4 out of 98 in level AAA, and 1 out of 42 in level AA web-safe colors, which satisfy WCAG 2.0 success criteria, become ineligible by aging.

Therefore, considering the effect of aging, while WCAG2.0 success criteria are mostly suitable for web-safe colors, we must be careful to use the colors on the boundary of the WCAG2.0 eligible area. And it is also recommended that we use 204 or more HSV value color against black background, because, if HSV value is less than 153, level AAA web-safe colors are not eligible for any age.

## KEYWORDS

Age-related Spectral Luminous Efficacy, WCAG 2.0, Luminance Contrast Ratio, Web-safe Colors, Color Universal Design

## 1. INTRODUCTION

Recently, web technology is widely used not only for the Internet, but also as application interfaces on many digital devices. And, toward the arrival of an aging society, a color scheme that considers the elderly attaches importance as color universal design. In the field of web, W3C consortium enacted the Web Content Accessibility Guidelines (WCAG) 2.0 as a standard of color universal design. Regarding the luminance contrast, WCAG2.0 success criteria (SC) 1.4.3 and 1.4.6 define the minimum luminance contrast ratios (7:1 [Level AAA], 4.5:1 [Level AA]). These required luminance contrasts are based on ISO-9241-3 (ANSI-HFES-100-1988), and the reduced visual acuity by aging. However, luminance used in these ratios is simply calculated by sRGB standard, without considering the decrease of spectral luminous efficacy by aging. Therefore, it is possible that some colors, which satisfy the WCAG2.0 SC but are inappropriate for the elderly, may exist.

In this research, we executed the precise measurement of spectral radiance for web-safe colors on LCD (sRGB mode). Next, we calculated age-related relative luminance and contrast based on age-related spectral luminous efficacy, which is defined in JIS (Japanese Industrial Standard) S 0031. Finally, we examine the validity of WCAG2.0 SC by comparing the theoretical sRGB luminance with the age-related luminance.

## 2. METHOD OF EXPERIMENT

### 2.1 Measurement of Spectral Radiance

We executed the precise measurement of spectral radiance for web-safe colors on LCD using spectroradiometer. Before measurement, we calibrate CIE xy chromaticity and display gamma of three source color lights to follow sRGB standard. All experiments were held in the dark room.

### 2.2 Calculation of Age-Related Relative Luminance

To derive age-related relative luminance and contrast that are comparable with WCAG 2.0 SC 1.4.3 and 1.4.6, we use age-related spectral luminous efficacy, defined in JIS (Japanese Industrial Standard) S 0031.

According to JIS S 0031, Age-related luminance  $L_{(a)}$  can be calculated by following formula.

$$L_{(a)} = \sum_{400}^{700} S_r(\lambda) V_{(a)}(\lambda)$$

$S_r(\lambda)$  : Spectral radiance       $V_{(a)}(\lambda)$  : Age-related spectral luminous efficacy

Normalizing  $L_{(a)}$  by white brightest light with standard spectral luminous efficacy, we can apply this ratio  $L_{n(a)}$  to the calculation of contrast and comparison with WCAG2.0.

$$L_{n(a)} = \frac{L_{(a)}}{\sum_{400}^{700} S_{r,white}(\lambda) V_{std}(\lambda)}$$

$S_r(\lambda)$  : Spectral radiance of brightest white light       $V_{std}(\lambda)$  : CIE standard spectral luminous efficacy

## 3. RESULTS

### 3.1 Comparison of Age-Related Luminance with WCAG2.0

To elucidate the effect of aging, we define the luminance ratio  $R_{wcag(a)}$  by following formula.

$$R_{wcag(a)} = \frac{L_{n(a)}}{L_{wcag}} \quad L_{wcag} : \text{Theoretical luminance calculated from the sRGB definition}$$

And we transform sRGB into HSV color system to see the hue distribution, because HSV is color appearance system and is widely used in the field of computer graphics. The result of calculation for 30 web-safe colors at maximum HSV saturation is shown in Figure 1, where vertical axis represents for Luminance Ratio  $R_{wcag(a)}$  and horizontal axis for HSV hues (12 degrees interval: R(0) corresponds to #f00 and M(300) to #f0f).

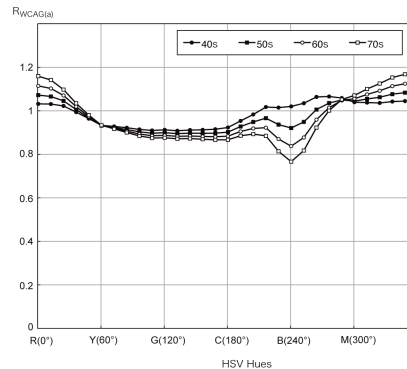


Figure 1. Luminance ratio against WCAG2.0 for 30 HSV colors at maximum HSV saturation

It turned clear that due to the decrease of spectral luminous efficiency by aging, age-related luminance also decreased especially at the bluish region centered on HSV hue 240 degrees.

### 3.2 Age-related Luminance Contrast

To verify the validity of WCAG2.0 SC, we calculate the age-related luminance contrast  $C_{(a)}$  that uses the same definition as in WCAG2.0.

$$C_{(a)} = \frac{L_{n(a),fg} + 0.05}{L_{n(a),bg} + 0.05}$$

In this formula,  $L_{n(a),fg}$  means the normalized age-related luminance of the foreground color and  $L_{n(a),bg}$  for the background.

The results of comparison  $C_{(a)}$  with WCAG2.0 are summarized in Figure 2. Figure 2 expresses HSV color planes (cone model). Each small square on these planes corresponds to web-safe color. The light gray areas indicate the set that satisfies WCAG2.0 SC 1.4.6 class AAA, and the dark gray areas stand for SC 1.4.3 class AA against black background.

As a result of calculation of age-related luminous contrast  $C_{(a)}$ , we have found out that 4 out of 98 in class AAA, and 1 out of 42 in class AA web-safe colors, which satisfy WCAG 2.0 success criteria, become ineligible by aging. These colors are pointed by black circle or rounded rectangle. We must be careful to use these colors whereas WCAG2.0 success criteria are mostly suitable for web-safe colors.

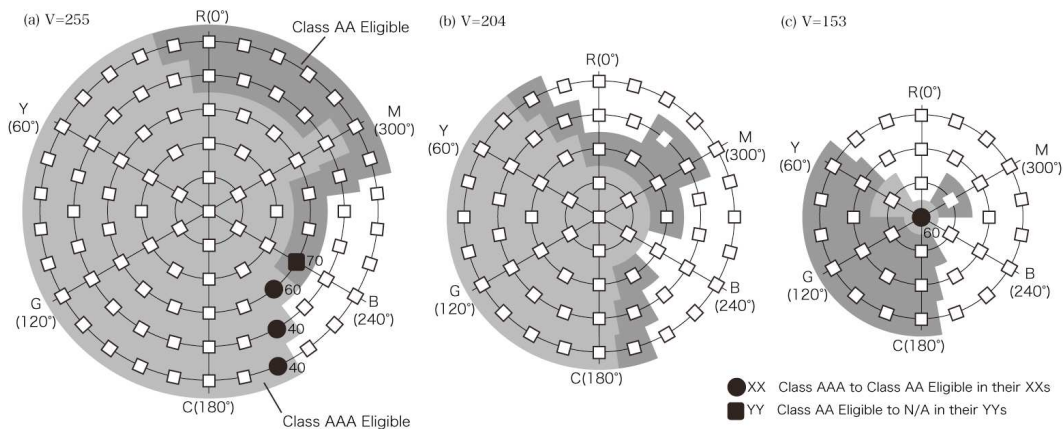


Figure 2. WCAG2.0 eligible areas and exceptions against black background at maximum HSV value.

## 4. CONCLUSION

To evaluate the validity of WCAG2.0 SC 1.4.3 and 1.4.6, we executed the precise measurement of spectral radiance for web-safe colors on LCD. Next, we calculated age-related relative luminance based on age-related spectral luminous efficiency. Finally, we verified WCAG2.0 by comparing the age-related luminance contrast with the theoretical sRGB contrast. As a result, reduction of luminance by aging is observed especially at the bluish region, and a number of colors that satisfy WCAG 2.0 SC, become ineligible by aging.

## REFERENCES

- Japanese Standards Association, 2004. *Guidelines for the elderly and people with disabilities -- Visual signs and displays -- Specification of age-related relative luminance and its use in assessment of light.*
- World Wide Web Consortium (W3C), 2008. *Web Content Accessibility Guidelines (WCAG) 2.0.*



# DEVELOPMENT OF RESIDENTIAL MONITORING SYSTEM BY MEASURING FIXTURES ACCELERATION

Macky Kato<sup>1</sup>, Yoshie Shimodaira<sup>1</sup> and Takeshi Sato<sup>2</sup>

<sup>1</sup>*Nagano Prefectural College, 49-7 Miwa 3 Nagano, 380-8525, JAPAN*

<sup>2</sup>*Jissen Women's University, 4-1-1 Osakaue, Hino, Tokyo 191-8510, JAPAN*

## ABSTRACT

Number of elderly people who live alone or elderly only couples are going to increase as ratio of elderly people increase in Japan as same as the other developed countries. They often have trouble with their health which needs care by someone. However most of their families live away from them. In this study we tried to develop remote monitoring system by measurement of fixture movement at home to observe the elderly people's daily behavior for safety. Through eight days experiment at rural area in Nagano, Japan, the acceleration measurement system recorded the door movement to observe the inhabitant's characteristics. Extracted characteristics from the results has shown the inhabitant's behavior which can be used for criteria of their activities. Elderly people who live alone or elderly only couples will be able to be monitored their safety by this automatic confirming system without violation of the inhabitant's privacy.

## KEYWORDS

Remote monitoring, Elderly people, Acceleration sensor, Fixture of house, Activity of Daily Life

## 1. INTRODUCTION

Many of elderly people especially in rural area live alone or live with just their partner. Because there are few jobs in the rural area to keep the young people's life, most of their children have to go out to the city area where they can get jobs. On the other hand, elderly people's physical depression can become the reason of accidents in their home such as falling accident. If the elderly people left in the rural area have troublesome when they stay home alone, they cannot help themselves. Thus they need some kinds of communication system in their own home to keep their safety and the family ties for "well being".

One of the methods to communicate remotely with someone is telephone. Nowadays, they can also communicate with video images through the Internet. Communication with family and friends by such devices can be considered as active relationship for the elderly people. On the other hand, security system to confirm the situation of home is needed to call someone at emergency. Because the elderly people do not need to operate the devices themselves, it can be considered as passive relationship. Conventional method to monitor the safety at home is monitoring camera, which records the inhabitant's all behavior all day. However too much monitoring makes them feel that they are under surveillance such as "1984" by George Orwell. Thus the current studies use non-camera method to confirm the elderly people. For example, there are confirmation systems by ultrasonic radar, GPS, and so on. Though, most of current methods need cost expensive, while analysis of frequency of using electric devices such as illumination is inexpensive. Switching time of illumination can be one of the signals of activity at home except the time when they do not need the illuminations. However, we have to confirm the elderly people's safety continuously. Developing of middle stance monitoring system is needed to solve the problems between safety and privacy.

In this study, we suggest the different method to confirm elderly people's daily behavior. To solve the privacy problem, we suggest observation of the inhabitant's behavior at home by not direct observation but by indirect observation of fixtures. That is to say, they have to use fixtures like doors at each point in the daily behavior. From this point of view, the measurement of the fixtures movement may become one of the indicators to show the inhabitant's behavior. The purpose of this study is verification of the daily behavior rhythm by measuring the movement of fixture in the house with using acceleration sensor. We practice the eight days monitoring experiment at an elderly couple's house by the acceleration sensor system to

investigate the behavior rhythm. The result will be able to become one of the criteria of elderly people's daily life, which may be applied to development of safety monitoring system without violation of privacy.

## 2. METHOD

Subject of this study is a couple who lives at farmer's house in Nakajo district, rural area of Nagano, Japan. Husband, a journalist of local weekly newspaper is 70 years old and his wife, housewife is 66 years old. They both are health. The developed system has been constructed from two parts, measurement system and remote network. The house has 6 rooms and many doors between each room. Through the interview with the subject, we selected the sliding door between living room and bed room as a fixture which has been used more frequently than the others, because the husband has a computer in his bedroom and he often use it.

Acceleration sensor AS-5GA (KYOWA, Tokyo) has been fixed on the edge of the sliding door (900mm width, 1,890mm height). The amplifier device EDX-100A (KYOWA, Tokyo) has recorded continuously the sensor output in its memory for eight days. Sampling frequency is 100 Hz. The amplifier device is controlled by software DCS-100A (KYOWA, Tokyo) in PC which connected to the Internet. PC has been connected to another PC through Virtual Private Network to show the real time movement of the door. Observers can see the variation of the acceleration of the door from other place and control the device remotely. After the eight days measurement, we would analyze the acceleration tendency to extract the rhythm of elderly couple's daily behavior from view point of short and long term. The tendency of the door movements would be able to be used as criteria of inhabitant's health.

## 3. RESULT

Figure 1 shows the typical acceleration wave of inhabitant's door opening and closing movement. Two types of acceleration have been observed in the figure, the first one is door opening and the second one is door closing. Acceleration time at the start is shorter than the deceleration time at the end of movement in both of the wave. At the end of the acceleration of door closing, a large and short wave occurs. All the waves through the experiment have characteristics of acceleration wave such as Figure 1. Figure 2 shows the acceleration wave in the range of 24 hours for the eight days. Several movements can be observed around morning and evening, however, the movements do not occur through midnights.

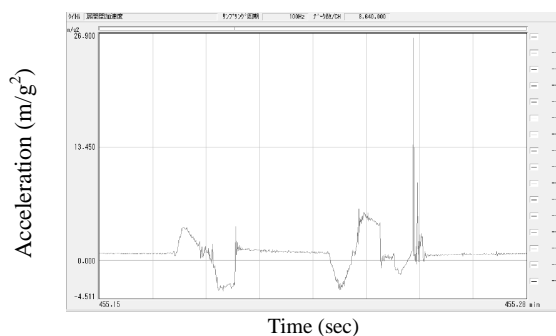


Figure 1. Acceleration wave of opening and closing door

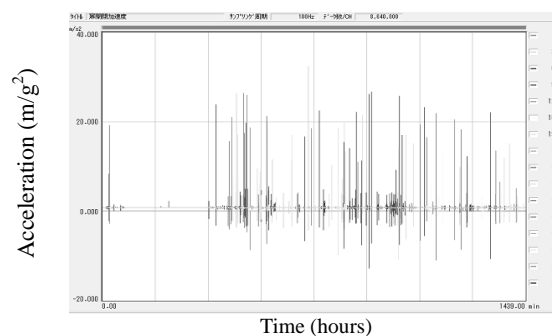


Figure 2. Acceleration wave in range of 24 hours through the eight days

The large and short acceleration at the end of shutting the door in Figure 1 always occur. The frequency of this kind of acceleration wave may be regarded as the volume of activity at home in a day, because the frequency of this wave means the frequency of opening and closing door. Figure 3 shows the histogram of frequency of such large acceleration movements through all the experiment days. Most frequent hours are around 9 and 17. The acceleration in day time from noon to 17 can be observed just in a few days.

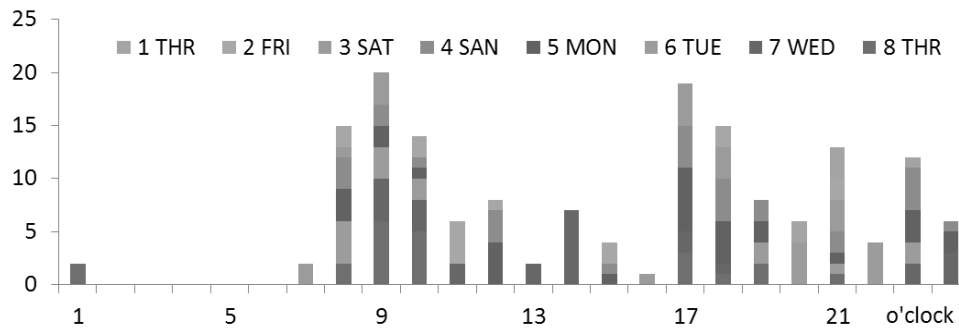


Figure 3. Frequency of large short acceleration in 24 hours through the eight days

## 4. DISCUSSION

Typical wave of door movement in Figure 1 has few differences among all the waves, because the movement is a kind of skill-based behavior. The subjects are not conscious when they move the door, except the large wave which means the impact of shutting the door which can notice them the closing movement by the sound. Short term acceleration variation in Figure 1 may become one of the criteria, because the door of bedroom in this case was used by only the inhabitant. It means that we do not need to consider of difference among individuals to monitor the inhabitants behavior.

Thus the wave of long term acceleration will be better than the short one to confirm the activity, because the wave in long term will have some variation and rhythm. Figure 2 shows the variation of acceleration wave. Frequency of door movement has some kinds of rhythm in 24 hours such as morning and evening with difference among each day. The reason why many of the impact waves occur in morning and evening would depend on the characteristics of the subjects. They are farmer in rural area and they keep regular hours.

Frequency of impact wave in Figure 3 also means the subjects' characteristics. Activities around 9 o'clock probably consist from morning preparing behavior such as breakfast, dressing and so on. Activities around 17 o'clock probably consist from two kinds of behavior, the wife's housekeeping and husband's computer operating. These are not the only one kind of behavior every day however they occur frequently.

This study is a kind of pilot study to investigate the rhythm of movement of fixture at home. If we use another door the wave would become different. It can be said that we will be able to use the all fixtures with sensors to confirm the safety more elaborately. When the observing trigger program finds out the largely different movement of door through monitoring, it might mean the inhabitant's trouble. Observer will be able to contact to the inhabitant and confirm their safety from other places. Such relationship between elderly people and observers will be able to keep eyes on them without violation of their privacy.

## ACKNOWLEDGEMENT

This study was supported by TOSTEM Construction Materials Industry Promotion Foundation and Japanese Grant-in-Aid for Scientific Research. We should like to express our thanks to the subjects.

## REFERENCES

Taro Sugihara, et. al., 2008, Effects of Cameras and Monitors on Caregivers' Work Stress in the Group Home, Institute of Electronics Information and Communication Engineering Technical Report, Japan, 107(555), pp57-62

# LEVERAGING OCR TECHNIQUE IN VIRTUAL KEYBOARD IMPLEMENTATION

Jure Mornar<sup>1</sup> and Andrina Granić<sup>2</sup>

<sup>1</sup>*Siemens d.d., Communications, Media and Technology, Put brodarice 6, 21000, Split, Croatia*

<sup>2</sup>*Department of Computer Science, Faculty of Science, University of Split, Nikole Tesle 12, 21000 Split, Croatia*

## ABSTRACT

Recent technology trends have led to a need for advanced HCI to provide a user with convenient input solution in the form of keyboard that would not require any desk space, be wireless and able to disappear when not in use. The aim of the paper is to present an easy way to improvise virtual keyboard that will require no hardware support besides the web camera attached to personal computer (PC) leveraging the method of Optical Character Recognition (OCR). Piece of paper with handwritten characters can be turned into virtual keyboard when webcam attached to the PC tracks finger movements and translate them into keystrokes in a device applying the OCR method. Touching handwritten characters on a white surface (paper) generates a unique electronic signal corresponding to the character that was touched. The implemented system proves valuable, but further research will be required to validate the findings of this study and to address identified drawbacks.

## KEYWORDS

Virtual keyboard, Optical Character Recognition, Human-Computer Interaction, Intelligent Interfaces.

## 1. INTRODUCTION

Nowadays, computing is not limited to desktops and laptops; it has found its way into mobile devices like palm tops and cell phones. What has not changed for the last 50 years is the input device, the good QWERTY keyboard. Alternatives came in the form of handwriting recognition, speech recognition, "ABCD" input (for SMS in cell phones) and the like. However, they all lack the accuracy and convenience of a standard keyboard. Speech input has an added issue of privacy. Even folded keyboards for personal digital assistants (PDAs) are yet to catch on. Thus a new generation of virtual input devices is now being paraded, which could drastically change the way we type. Consequently, last few years have brought a lot of related work and solutions that can provide new and innovative ways for text input supported by various hardware components such as touch screens, tethered gloves, lasers, infrared cameras and so on. Although some of these solutions are already in use for commercial purposes, the fact is that they rely on expensive hardware support. This paper brings the idea of rather easy implementation of the virtual keyboard using just web camera (that most laptops and desktop computers already have) and applying the method of optical character recognition.

## 2. BACKGROUND TO THE RESEARCH AND RELATED WORK

Although new human-computer interfaces have been implemented to provide multiform interactions between users and machines, basis for most human-to-computer interactions remains the binomial keyboard/mouse. Ordinary keyboards however, to be comfortable and effective, must be reasonably sized. Thus they are cumbersome to carry and often require wiring. Concept of virtual keyboard was introduced to overcome these problems. A large body of related work exists for physical keyboards and typewriters, their characteristics, usability, efficiency, history, and anthropomorphic backgrounds [Noyes, 1983; Dvorak, 1943; Alden, Daniels and Kanarick, 1972]. As the keyboard becomes one of two largest components of computing devices (next to the display), research on smaller and more mobile text entry methods and devices has made great strides [MacKenzie and Soukoreff, 2002]. Finger tracking and finger tracking based interfaces have

been an actively researched problem for several years now. For example, glove-based systems, such as the “Key-glove” [Won et al, 2001], require the user to wear a tethered glove to recognize signal variations caused by the movement of fingers. Senseboard for example has developed a virtual keyboard system based on two devices made of a combination of rubber and plastic. The devices recognize typing events by analyzing the data from pressure sensors attached to the user’s palm. Computer vision researchers as well have made significant advances in the development of vision based devices that require no wearable hardware. Canesta and VKB [VKB Ltd] for example have designed virtual keyboard systems using infrared cameras to detect the interaction between the fingers and a projected image from a laser diode.

All those techniques and approaches for implementing virtual keyboards have something in common – they are quite expensive to implement as they require hardware support in the form of various touch screens, infrared cameras, laser diodes, optical or electronic beams, tethered gloves and so on.

### 3. PROPOSED SYSTEM PHYSICAL SETUP

Web camera is placed above the input surface, with an angle of approximately 30°-60° (in order for OCR method to perform well) facing the working area. Resolution and distance of the web camera defines size of the working area. This working area should be comparable to a full-size computer keyboard. Inside the working area there is a piece of white paper with hand-written letters on it (“ABCD” in this case). With the web camera facing this piece of paper, we can turn it into a virtual keyboard. System physical setup is shown on Figure 1.



Figure 1. Proposed system setup

#### 3.1 Image Processing and Optical Character Recognition

After the web camera is connected, program can start with process of grabbing and processing images. Thus captured image is inverted, so it has to be rotated. In addition, image will be converted to gray-scale and thresholded to be prepared for applying OCR method. Process of image processing is shown on the Figure 2.

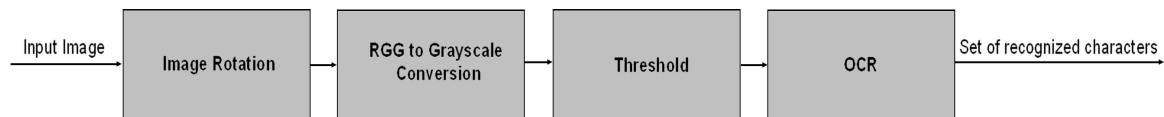


Figure 2. Image processing steps

First, the size, width and height of the image are noted. The transformation of the image is performed by dividing the width and height of the image into halves and pixels are grabbed from the image. According to the specified angle the inverted image is rotated and it is converted into normal image. Since operations are performed only on grey scale images, the grabbed colour images are converted into grey scale image and thresholded to make optical character recognition method much more effective. After grabbing, rotating, transforming to black and white and thresholding frame is ready for performing Optical Character Recognition (OCR) method. OCR method is performed using previously Google Tesseract OCR API. Processed image of the keyboard is passed to the recognizeText() method of the OCR class instance which returns String with recognized characters from the input image (picture of keyboard).

#### 3.2 Finger Movement Tracking

Finger movements can be tracked by continuously grabbing frames from web camera and analyzing the set of characters recognized using OCR method. If the set of recognized characters was “ABCD” for one frame, and after processing next frame set is like “ACD” (character “B” is missing) it can be assumed that character “B” on the paper is covered by the finger (character “B” is clicked). However, when using this approach we

must be careful to avoid undesired outcomes, like, for example, situation when we try to write “AC” and frame is grabbed just in the moment when our finger is above the character “B” may result in wrong conclusion that character “B” is clicked. This problem is solved by introducing a logic which will assume specific character as “pressed” after N consecutive frames shows that this character is missing from set of recognized characters. This number N depends on web camera grabbing frequency and together with the “M” parameter (idle time after performing OCR method in milliseconds) defines “virtual keyboard sensitivity”. Since the keys “are not real” and conclusion on whether the key is pressed is made solely on the basis of OCR method results, a mistake can occur in a way of “unwanted multiple clicks” if keyboard is too sensitive (when N and M are too low). When it detects that “key is pressed”, algorithm plays the notification sound similar to sound produced by real keyboard keystroke to provide user with a good keyboard feedback which can help avoiding the problem described above. Aforementioned algorithm is presented by pseudo-code in the Figure 3.

```

While program is running
Perform OCR method on grabbed picture
  If OCR method result says that some character is clicked
    If clicked character is equal to the last clicked character
      Increase temp character counter
      If temp character counter is equal to N
        Assume character as clicked
        Reset temp character counter to zero
    Else
      Set last clicked character to the value of clicked character
      Set temp character counter to one
Wait M milliseconds

```

Figure 3. Finger movement tracking algorithm presented by pseudo-code

#### 4. DEVELOPED SYSTEM REVIEW

Web camera was used during the development of application as a device for capturing images of keyboard and fingers. Keyboard layout is determined by the layout of letters (characters) on the surface of white paper. In this paper, the keyboard consists of four hand-written letters (ABCD).

The graphical user interface (GUI) of the developed system can be divided into four parts, as shown in the Figure 4. Part of the application numbered 1 presents view from the webcam. There we can keep track of all activities related to the virtual keyboard, movements of the fingers, etc. Webcam should be suitable so that strictly focuses on the surface of paper with characters written on it. Number 2 indicates the area of three buttons. Button *Grab* is used to capture the initial image of the keyboard using web camera. So captured image will be displayed in the area numbered 4 (a more detailed description follows). Pressing *Scan* button initialize process of image processing and optical character recognition on the captured image. The result of image analysis, i.e., a set of identified characters is printed in the application area marked with the number 3. Number 4 indicates the multifunctional part of application GUI. In the initial phase this area contains a text with explanations and instructions while later it is used as area containing input from the virtual keyboard.

When using developed application, first step is to connect the web camera and set it so that strictly covers the surface of paper with letters written on it. Once the webcam properly set, pressing the button *Grab* grabs the picture from a webcam. So captured image is displayed in the lower part of the application GUI and optical character recognition method is performed on grabbed image when button *Scan* is pressed. Characters recognized from image are printed out in the part of application GUI called *Recognized characters*. If the characters written on paper are not successfully recognized, most probably letters are written incomprehensible or lighting in the room is not adequate. When all the characters from the paper are successfully recognized, by pressing the button *Start* virtual keyboard starts monitoring finger movements. Input from the virtual keyboard is shown in the multifunctional part of application GUI, as shown in Figure 5.

#### 5. CONCLUSION

The ultimate aim of this paper was to present the idea of how to implement a virtual keyboard leveraging the method of optical character recognition (OCR) and using webcam as the only required hardware device. Implemented system provides users the ability of using a piece of paper with hand-written letters as virtual keyboard with good performance and ease of use. Furthermore, users have a freedom to determine the

position and shape of virtual keyboard, as well as the size of written letters as long as they are not too small for OCR engine to recognize them and not too large to be out of camera's field of view.

However, there are few drawbacks related to the implemented system. Although considered to be the best and most accurate open source OCR engine, the biggest shortcoming of the developed system lies in the imperfection of Tesseract OCR API (reflected in the inability to recognize characters in conditions of inadequate lighting). Furthermore, the implemented algorithm for monitoring finger movements and recognition of input from virtual keyboard currently supports only one form in which keyboard characters are printed in one line. Aforementioned algorithm can be easily extended to support keyboard layout where characters are printed in multiple rows. Additionally, this work can be expanded using more complex hardware support like laser technology to draw the keyboard instead of writing characters on piece of paper, or using powerful 3D camera which will be able to take finger-paper distance together with OCR results into the account while making decision about clicked characters. The obvious next step will be a development of an evaluation strategy for implemented system assessment which result/feedback will be a valuable guide for redesign and further improvements of the virtual keyboard system. Accordingly, additional thorough research will be required in order to validate the findings of the conducted study.

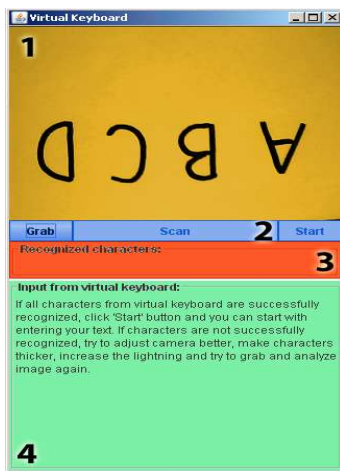


Figure 4. Developed system GUI



Figure 5. Monitoring for user's input

## ACKNOWLEDGEMENT

This work has been carried out within project 177-0361994-1998 *Usability and Adaptivity of Interfaces for Intelligent Authoring Shells* funded by the Ministry of Science and Technology of the Republic of Croatia.

## REFERENCES

- Noyes J. 1983. *The QWERTY Keyboard: A Review*. International Journal of Man-Machine Studies, Vol. 18, pp. 265-281.
- Dvorak A. 1943. *There is a Better Typewriter Keyboard*. National Business Education Quarterly, Vol. 11, pp. 51-58.
- Alden D.G, Daniels R.W. and Kanarick A.F. 1972. *Keyboard Design and Operation: A Review of the Major Issues*. The Journal of the Human Factors Society, Vol. 14, No. 4, pp. 275-293.
- MacKenzie I.S. and Soukoreff R.W. 2002. *Text Entry for Mobile Computing: Models and Methods, Theory and Practice*. Human-Computer Interaction.
- Won D. H, Lee H. G, Kim J. Y. and Park J. H, 2001. *Development of A Wearable Input Device Recognizing Human Hand and Finger Motions as A New Mobile Input Device*. International Conference on Control, Automation and System, pp. 1088-1091.
- VKB Ltd. (Available on <http://vkb.co.il>, accessed on 25.04.2011)

# DESIGN OF A FOLDING WORKBENCH AND A VIRTUAL REALITY SIMULATION SYSTEM USING IT

Hyun Suk Lee, Bon-Wook Gu, Jongwon Back and Tae-Jeong Jang  
*Electronic and Communication Engineering, Kangwon National University  
Hyoja2-dong, ChunCheon, 200-701 South Korea*

## ABSTRACT

In this paper, a virtual reality simulation system is presented. That simulation system uses a workbench system and a haptic device. A user can see stereoscopic objects through workbench, and he or she can feel tactile information like real using the haptic device. And we designed workbench system which can be folded for easy movement and simple storage. Also we made stereoscopic images to intensify simulation reality by actual measurement.

## KEYWORDS

Virtual reality, Workbench, Haptics, Haptic device, Haptic interface, Stereoscopic system, Training simulation system

## 1. INTRODUCTION

Virtual reality system has been used extensively in several areas like sciences, medical systems, education, and entertainment. However, if interaction environment dose not satisfy a user enough, the effect of virtual reality could be lowered. Because a person gets information about the surrounding environment through the interaction of the senses as touch, sight, hearing, etc. Therefore, haptic interface is used on the virtual reality simulation system. Haptic interface is the system which makes interactions with user by sense of touch. It makes a user feel as real while touching virtual objects by haptic device, primarily.

A general haptic interface simulation system can be effective to express reaction because of using force feedback. But, since actual workspace and a virtual work space does not match, users face a difficulty they should adapt to the new simulation environment definitely. To overcome these problems, the system uses the workbench system for matching monitor and real workspace. Workbench system can make a virtual reality system more immersive.

In this paper, we built a virtual reality training simulation system using a folding workbench system, stereoscopic objects and a commercial haptic device (PHANToM Omni). And the folding workbench system is designed for mobility and easy storage.

## 2. THE SYSTEM DESIGN AND IMPLEMENTATION

### 2.1 The System Configuration

The workbench system configuration we use for this simulation system is shown in the Fig 1. This workbench system includes a stereoscopic display and a haptic device. The stereoscopic virtual objects on the monitor are shown to a user by reflection of the mirror. A main computer is for controlling this monitor and a haptic device, and an extra monitor is for user's control panel. Menus for functions and/or status of the overall system are shown on the control monitor.



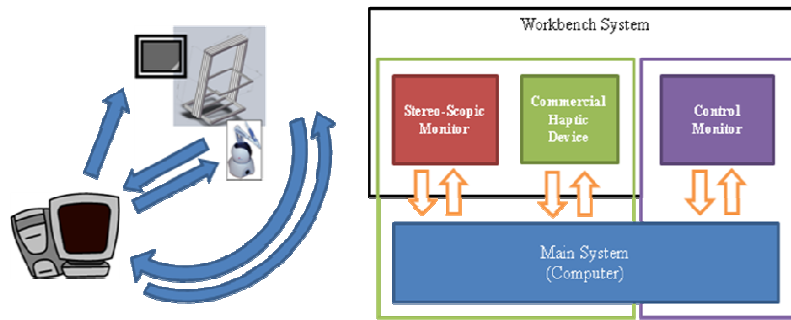


Figure 1. The workbench system configuration for a training simulation system

## 2.2 The Device Design and Implementation for Matching 3D Virtual Space and Real Space

Workbench System is the immersive system made for matching a virtual space and the real work space. Usually a haptic simulation system is inconvenient because the virtual space a user can see and the real work space did not match. Controlling depth of stereoscopic objects helps a user to feel like real work while working on the system.

### 2.2.1 The Device Design and Implementation

When designing a workbench for virtual reality simulation system, there are several considerations. In case of the common workbench, the tilted angle of virtual object on the reflector is the same as the tilted angle of the monitor. Since tilted virtual objects might be less immersive, we have to consider parallel positioning between virtual objects and the surface of work. That is the first consideration. Second, the existing workbench system is usually fixed on one place since it is big and heavy shape. When you need to move the workbench from one place to another or when it is not in use for a long time, the large volume of the system is inconvenient. Training systems for education sometimes need high mobility and small volume since there might be various kind of such systems for students to learn and there might not be enough space to place all these systems. Therefore, the design of folding workbench is useful for such purposes- high mobility and easy storage.

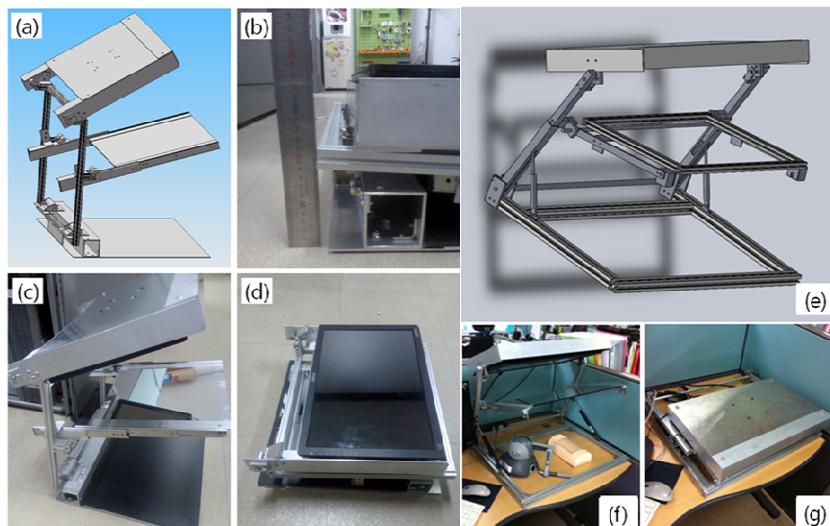


Figure 2. First workbench system (left a-d), edited system (right e-g)

The left side of Fig 2 shows 4 photos of the first workbench we made. This workbench is designed that all the loads are concentrated in the pole frame. Therefore, when we stretch the workbench, the elasticity of hinges consisting of spring-damper supports the heavy weight by increasing. These hinges make reaction for maintaining the stretched state. These facts make a user feel inconvenient when the user try to fold the system. Also, the workbench joint for folding is not stable enough because of heavy weight and excessive force by user. When a user tries to fold the workbench, the monitor faces the risk of breakage because the stereoscopic monitor is exposed to the outside. The 2<sup>nd</sup> workbench was created to overcome these limitations and we can see on the 3 pictures in the right side of Fig 2.

### 2.2.2 The Stereoscopic Image Rendering for Matching Virtual and Real Space

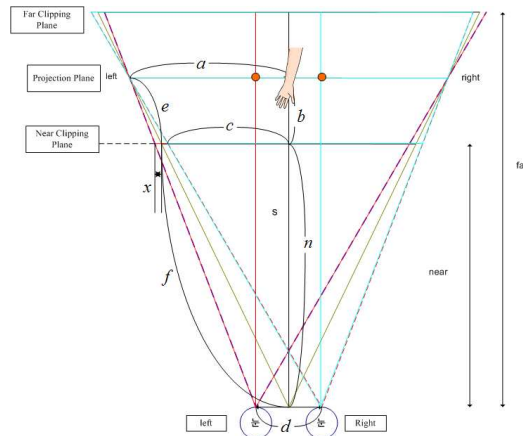


Figure 3. Axis asymmetric frustum perspective projection

From the workbench system, the method of rendering stereoscopic images is as follows. First of all, we assume user's eyes are located in a certain point. Second, we only consider that the projection plane has only horizontal parallax between left eye and right eye. Vertical parallax of both eye is assumed 0. For the projection plane, plane for the left eye is the same as the right plane. Usually, the a center of an important object is located on the projection plane. And far clipping plane is set as ground. In this condition, near and far clipping planes of each eye have difference. The  $x$  for the near clipping planes in Fig 3 is one of examples about a plane difference of both eyes. That difference makes a point difference of object, and it is the reason we can see a stereoscopic image using the workbench.

This method shows same features with the vergence control methods of the parallel stereoscopic camera when we consider the rendering result of stereoscopic images. It is known as asymmetric frustum parallel axis projection stereo (J. D. Brederson et al. 2000). We rendered the projection plane of stereoscopic images which has the same scale as a real work space of workbench by measurement. To confirm the results, we rendered stereoscopic images by placing some sphere objects in various positions. The results are shown in Figure 4.

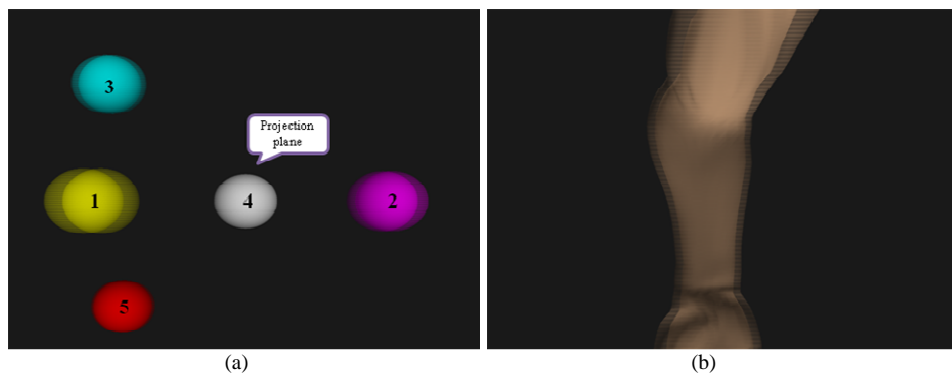


Figure 4. The result of stereoscopic rendering images

Images in Fig. 4 are checked as a result of stereoscopic images on the non-stereoscopic monitor. Fig. 4(a) is numbered according to the ascending order by the short distance from the near clipping plane. Sphere No.4 is located on the projection plane. There is no difference of binocular parallax in the sphere located on the projection plane. Fig. 4(b) is the rendered image of an arm model which is located on the projection plane by the same way. Each part of the arm model has a different position. Since the center of arm model is located on near the projection plane, the difference between plane for left and right eye is small. This method has the advantage of low image distortion than creating stereoscopic image method by crossing the focus and the axis.

### 3. CONCLUSION

In this paper, we implemented virtual reality simulation system by using a workbench system and a commercial haptic device. We can feel intuitive and immersive virtual reality in the system by matching virtual space and real work space. In virtual reality system, immersive system is an important factor for users. So we adjusted user's real work space and virtual space physically in the system and developed the stereoscopic by axis projection stereo. In particular, the folding workbench system architecture we proposed is easy to move and store as small volume. Currently, we developed the simulation system by using a commercial haptic device. In the Future, we will produce our own haptic device which will be specialized in the specific training that can show the real and the same reaction. That would be a good chance to develop simulation training applications suitable for the training. After that the system is expected to perform tests on.

### ACKNOWLEDGEMENTS

This research was financially supported by the Ministry of Education, Science Technology (MEST) and Korea Institute for Advancement of Technology(KIAT) through the Human Resource Training Project for Regional Innovation

### REFERENCES

Book

Sensible, OpenHaptics programming guide.

Journal

Daniel C. Cliburn, 2004. Virtual reality for small colleges. *Journal of Computation Sciences in Colleges archive*, Vol.19, Issue 4

Maupu, D. et al, 2005. 3D stereo interactive medical visualization. *Computer Graphics and Applications IEEE*, Vol.25 Issue 5, pp 67-71.

K.C.Kwon, et al. 2002. Automatic control of horizontal-moving stereoscopic camera by disparity compensation, *Journal of OSK*, vol.6, no.4, pp.150-155

Conference paper or contributed volume

Gu, Bon-Wook. et al. 2010. Workbench System for Virtual Reality Simulation. *conference of the Korea Society of Broadcast Engineers*, pp. 306-308.

Srinivasan, M.A. and Basdogan, C., 1997, *Haptics* in virtual environments: taxonomy, research status and challenges. *Comp. Graphics*, Vol.21, pp.393-404.

J. D. Brederson, et al. 2000. The visual haptic workbench. In *Proc. Fifth PHANToM Users Group Workshop (PUG)*.

R. Maciejewski, C. Seungmoon, D. S. Ebert, and H. Z. Tan. 2005. Multimodal perceptualization of volumetric data and its application to molecular docking. In *Proceedings of the 2005 World Haptics Conference (WHC05)*, pp. 511-514.

# A STUDY ON CATEGORY TYPE OF ONLINE SHOPPING MALL

Kristen Koeun Nam, Kyle Kim and Sungmin Hong  
*Hongik University*

## ABSTRACT

Form of fixed category of online open market site has deviation in their usefulness depending on the type of information search of users. This study is to look into the preferences of changeable categories which complement the demerits of these fixed categories and directly connected with the experiences of users in the real world.

As a result of measuring the preferences of changeable categories by eye-tracking test and in-depth interview, the result values are different depending on the types of online information search and shopping of users. As a result of eye-tracking, changeable category shows high fixation duration and proceeds in more concentrative manner than fixed category. This result can be said to have affirmative influences on the action of search. In addition, the type of non-purpose oriented shopping showed affirmative opinion such as menu attention and increase in convenience of changeable category, but obtained the result that users who have distinct goods to buy have strong tendency to use search rather than category classification which they recognize through their experiences. This study reckoned that changeable category would be useful as a ring that connects real world of users with a grade of ranks of online category in case phrases that can predict connection link of goods, elements of arousing interest in connection with hot trends, social communication function and offering of benefits which included.

## KEYWORD

Online open market site, Fixed Category, Changeable Category

## 1. INTRODUCTION

Online open market sites take the form of fixed category. The 'fixed' mentioned here can be standard classification system of top-down method which online open markets basically use, as form of category is the form of taxonomy. Though the form of fixed category is used in many internet shopping sites, users are going through many trials and errors before they find and buy the product they for search through categories. The reason why users cannot easily find the product is because the mental models of individual users are different.

Users are going to search for information (products) in manners different from each other. Fixed category has insufficiency in covering the types of various information searches of users. Therefore, the type of category which supplements demerits of fixed category online is necessary and we are to understand that the type should be form of category connected directly with the experiences of users. In addition, we have the purpose to look into the improvement points of category newly suggested through this category.

## 2. BACKGROUND

Moe (2003) classified the type of shopping into directed buying, search and deliberation, hedonic browsing and knowledge building depending on the method of search and period of purchase and maintained that each of the types has characteristics of search type. She described a distinctive characteristic of directed-buying visits is the shopper's tendency to exhibit very focused search patterns, indicative of the goal-directed motivation of the shopper. As a result, more product-level pages rather than category-level pages are viewed, as category pages provide a broader level of information and product pages provide more targeted and detailed information. Specifically, shoppers will likely view pages within a limited number of products and categories.

This study suggests changeable category which compensated this demerit of fixed category and expects that the grade of rank of category related to his experience will have affirmative influence on finding information (product). Changeable category is defined as the grade of rank of category as sales classification commonly used in real world (For instance a customer who is to buy electric bulb cannot easily recognize category including electric bulb, but electric bulb easily reminds him of electrical shop commonly used in the real world.) hence though they are not the products in the same kind, includes the category where products of diverse product families gather in one theme. Therefore, the research question of this study is:

**Q: Does Changeable category have affirmative influence on user finding information (product)?**

### **3. METHOD**

To verify the hypothesis, this study investigated preferences of fixed category and changeable category using in-depth interview and eye-tracking equipment with the object of 20 persons who use online shopping mall. For experiment objects 20 persons, we recruited the users of Korean open market 11st, and in-depth interviews including eyes-tracking experiments were progressed over a week and about 1 hour of time per person was spent.

### **4. RESULT**

By the result of an Eye-tracking Method, we were able to detect significant differences depending on the type of Categories. Comparing both Fixed and Changeable categories, changeable category type turned out to be stayed longer than the fixed category menu which examining in Heat map (see. Figure.1).

Based on the results of AOI, the value of existing fixed category is relatively high compared to the variable categories. For the case of changeable category, user wants to buy products associated with the product which can be found together. The various categories in one-depth without moving to the other categories can be recommended which seems to be showing positive results. The result of frequency is an important characteristic that pulls the attention of users in online shopping site. But if changeable categories are recognized as an ad, it will bring a negative feedback of category. By this result, attention is not always helps Usability.

By in-depth interview we could confirm that users' category preferences are different depending on shopping types of Moe(2003). The users of Search/Deliberation type showed affirmative opinion on menu attention, rise of convenience of changeable category.

With changeable category, users can more easily recognize the grade of rank of category which can be associated with experiences of users than with fixed category whose letters were small and which was complicated so it is reckoned to be accessed comparatively easier.

When using fixed category the movement to too many detailed categories was unsatisfied but changeable category comparatively has less movement to detailed categories.

The search using changeable category is a help for people when they use it to purchase product family vaguely.

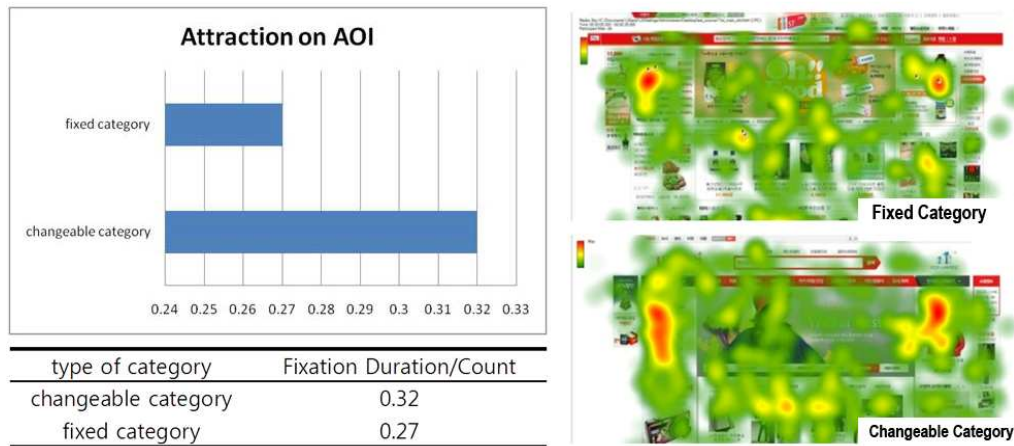


Figure 1. Attraction on AOI (all fixations are summed and the number is divided by the total fixation) & Heat map

On the other hand, the type of directed buying which used product level rather than existing category level showed negative opinion on changeable category, because the users who have strong tendency to use search method rather than category classification, which they recognize through their experience.

## 5. CONCLUSION

This study is to understand the preference of changeable category form which compensated the demerits that fixed category has. Changeable category as the grade of rank of category helps users to search online information through the recollection of experiencing in real world since the users usually have weak knowledge of searching in categories.

The limitation of this study has no procedure of measuring usefulness and preferences of fixed and changeable categories following diverse online information search types. Therefore, studies in the future also should confirm usefulness of changeable category by conducting measurement suitable for types after prior verification of online shopping type of users.

## REFERENCE

- Moe, W.W. (2003). BUYING, SEARCHING, BROWSING: Differentiating between online shoppers using in-store navigational click stream. *Journal of Consumer Psychology*, 13(1&2), pp. 29-39.

# NONVERBAL COMMUNICATION TOOL VIA USE OF PERCUSSION PERFORMANCES AND ITS EFFECTIVENESS

Yuji Kawase<sup>1</sup> and Yoshio Nakatani<sup>2</sup>  
<sup>1</sup>*Graduate School of Science and Engineering*  
<sup>2</sup>*College of Information Science and Engineering*  
*Ritsumeikan University*  
*1-1-1 Noji-higashi, Kusatsu, Shiga, Japan*

## ABSTRACT

In general people communicate their feelings to others via language. In some cases, however, nonverbal communication can be more suitable for conveying their feelings than verbal communication. This research proposes a tool that be used in communication via the sound produced by percussion instruments when people find it difficult to verbally express their feelings. Some cognitive psychology research supports the idea that feelings can be expressed via rhythms. With the method proposed here users can select the color that best matches their current feeling from a palette, and cases where the current feeling was best conveyed in the past are then searched and proposed. The system can be improved upon via use of the color palette. The effectiveness of the system was then verified in experiments, which revealed that users were able to choose an appropriate percussion performance that matched the way they felt. Moreover, people listening to the performance were able to understand the user's feeling. This system was thus proved capable of revealing the effectiveness of nonverbal communication made via percussion performance.

## KEYWORDS

Nonverbal communication, emotional communication, music, percussion

## 1. INTRODUCTION

Communication is essential in the daily lives of humans. In general people communicate their feelings to others via language. However, nonverbal communication can sometimes be more suitable for conveying feelings or thoughts than verbal communication. Congenital language disorders also do not pose a problem to communication. This then means that the common feelings to all can be conceived [1]. The need also therefore exists for non-verbal types of communication.

One of the issues in this study is supporting the needs of people who wish to convey their feelings non-verbally using the computer. A variety of methods of communicating non-verbally exist that include as gestures, facial expressions, painting, music, and so on. If it can be assumed that some people cannot verbally describe their feelings or situations it stands to reason that the need exists for them to be able to convey their feelings to others via use of other media. Moreover, it would also possibly be of support of any person who cannot express themselves very well.

This research proposes a tool that can be used to communicate via the sound of a percussion instrument what people may feel difficult to verbally express. Sounds produced by percussion instrument are linked to the feelings of users, thus resulting in a new type of communication tool. After the relationship between SOUNDS and FEELINGS are clarified, however, the mechanism in which SOUNDS and FEELINGS can be linked together is needed.

## 2. METHOD OF SELECTIONG SOUND FROM COLOR

The present study aims at communicating feelings using a percussion instrument, being based on the above findings, ensuring the user's feelings are adequately expressed, and identifying a method of confirming both the aforementioned [1][2][3]. The user's feelings are specified and music from a data base that expresses the feelings best then retrieved, which is then presented to the user. The user then transmits the best fitting performance to the other party.

Feelings were first specified using color. This is actually used in daily life, and it has been widely acknowledged that different colors provide certain meanings to people at both the cultural and emotional levels [4]. A method of selecting two colors to express a user's feelings at that time was used due to people's feelings tending to be very complex. People generally do not feel only "A little more lonely than happy", "Angry but slowly feeling better", etc. or any one particular feeling.

A ratio is then applied to the two selected colors. For example "Red is 9 and blue is 1," and the sound source then extracted using that ratio. Colors and percussion instrument performances are located in the same "sensibility space" calculated by using the semantic differential method and the factor analysis. A sensibility score was then calculated using the two selected colors. An input ratio is used to divide the distance of the two colors. The nearest percussion instrument performances to the divided position is selected and proposed to the user.

Moreover, this research also takes user's individual variations into consideration as the results obtained from the questionnaire cannot be asserted to positively identify user's feelings, although nor a gap. Moreover, it is insignificant, even if a system that takes a user's individual variations into consideration is recommended. If a sound obtained from a color being selected or an adjective approximates the user's feeling it closes in on the position the user showed. The user can then customize it to the sound, if not suitably opposite. The method of customization is carried out using same content of sound as the extraction, with the position shown then becoming the new position of the sound. Naturally the sound source is updated with that factor. A concrete image is provided in Figure 1.

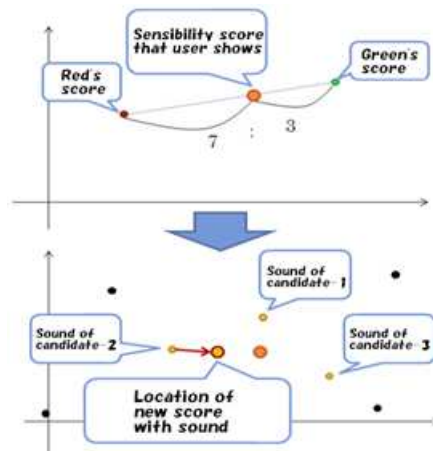


Figure 1. Image of proposal

There is a possibility that a gap between the user's personal feelings can result because the factor score of the music and the color is based on the result of a questionnaire provided to only two or more testers. If the sound obtained from selecting a color roughly suits the user's feelings the position and factor score for the user are brought closer, and updated by the music. A system that appropriately the extracts feelings that each user wishes to express using music can thus be expected by repeating this method.

## 3. EVALUATION

The expectation is that this method can be used express feelings in detail and accurately very well, even in the case of the user understanding their own feelings, from the feature included in the proposed method,



along with non-verbal expressions. Moreover, it will become easier to select the music from user's inputs because the sensibility score to be selected is updated when expressing the music. The following two hypotheses were set for use in experiments.

1. The listener can actually understand user's feelings from the tune the user chose.
2. Whenever the system is gradually used music matching the user's feelings is expressed.

An evaluation for the sake of comparison took place a total of two times after it had used it for one week from when the music database had been initialized. It treats as 2 and 6 evaluations with the values of 1 and 7, respectively.

The first hypothesis was then verified. The question "Did you feel that the expressed music applied to your current feelings?" was put to the user. Moreover, the questionnaire was made available to the listener at the same time, "Were the user's feelings imaginable via listening to the music?" The listener's evaluations were comparatively high, with the evaluation exceeding 4 in the results being 75%. However, the listener resulted in high appraisals opposing those seen when evaluating music in that it did not match the current feelings of a certain user.

The second hypothesis was finally verified. It was similarly evaluated using a question the user was presented with in the questionnaire, "Was the music that agreed with the feelings after it had been used before expressed?". It can be seen that music which expressed the feelings of user, at 67% or 5.0 on average, as in the results, resulted from a one week long case. However, one person did evaluate it quite low along with one that evaluated it as not having changed very much, too.

## 4. CONCLUSION

This paper proposed a sensibility communication system utilizing percussion instrument performances, which it appears to have achieved. The system is composed of two parts: a "Search engine" and a "Sensibility score updater". The colors are then assumed to be part of the decision making involved in feelings, be related to music, and expressible. Moreover, the expressed music uses a "Sensibility score updater" to take individual variations into consideration, thus making the score close to where the user wants to express the sensibility score of the music used. Individual variations are thereby taken into consideration using this technique.

Listeners were also able to take the user's feelings into consideration. Moreover, it became easier for the music used for the user's feelings to be expressed by updating the sensibility score of that music, with the effectiveness of the appropriate method of that having been revealed above. However, all the testers were not as successful. There were gaps in what the user's feelings actually were and feelings the music invoked in listeners.

## REFERENCES

- [1] Toshihiko ENDOU, 1996. *Origin of emotion—Theory of evolution of affect·culture theory*, Iwanami Publishers, Tokyo, Japan. (in Japanese)
- [2] Shinobu IKOMA, Tadashi KIKUCHI, 2006-07. Emotional communication by improvisation with percussive timbre. *The institute of electronics, information and communication engineers*, IEICE Technical Report HCS2006-26(2006-07), pp.19-22. (in Japanese)
- [3] Akio YAMASAKI, 2006. Communications of feelings through performance by professional drummer. *Osaka Shoin Women's University, The Human Science Research Bulletin 2006*, No.5, 111-122 (in Japanese)
- [4] Tomotsugu KATAOKA, Toshie NAKAMURA, Chika NAGAOKA, Satoshi KAWASE and Maria Draguna, 2003. Expression of Emotions with drumming performance. *Human Interface 2003*, Tokyo, Japan, pp.449-452 (in Japanese)
- [5] Saeko INOUE, 2001. Consideration concerning effect of color therapy, Osaka University of Economics, *Graduation Thesis in 2001* (in Japanese)
- [6] Makiko Okita, Yuki Nakaura, and Hidetsugo Suto, 2009. Nonverbal Communication System Using Pictograms. *HCI international 2009*, San Diego, CA, USA, pp.720-724
- [7] Marak L. Knapp, Judith A. Hall, 2009. *Nonverbal Communication in Human Interaction*. Lyn Uhl, Boston, USA

# VR INTERACTION TOOLS FOR MOTOR IMPAIRMENT SIMULATION

Marco Fontana and Marcello Carrozzino  
*PERCRO Laboratory, Scuola Superiore Sant'Anna, Pisa, ITALY*

## ABSTRACT

The work that is presented here deals with the preliminary developments of a set of Virtual Reality (VR) tools intended for the simulation of motor impairment. This work is developed in the framework of the European Project VERITAS whose objective is to develop a set of tools to support inclusive design of products and services. Among the main objectives of the project it is planned to develop of an Immersive Simulation Platform (ISP) based on virtual reality technologies where designers assess the result of their design work assuming the point of view of the disabled with the aim of allowing her/him to directly experience a specific disability. To this aim a set of custom Interaction Tools (IT) have been developed. In this paper we introduce the Interaction Tools (ITs) for the simulation of motor disabilities that make use of newly developed interfaces. In particular we present three interfaces dedicated to the simulation of Kinematics, Dynamics and Control functional limitations.

## KEYWORDS

Inclusive Design, Disability Simulation, Virtual Reality

## 1. VERITAS INTERACTION TOOLS

The European Project VERITAS (Virtual and Augmented Environments and Realistic User Interactions To achieve Embedded Accessibility DesignS) aims at developing a set of tools to support designers for accessible design of products and services in the domain of automotive, smart living spaces, workplace design, infotainment and personal healthcare and wellbeing. These tools include Virtual User Models that integrates Physical, Cognitive and Behavioral/Psychological models and simulation platforms. A further scope of the project is to develop of a virtual reality Immersive Simulation Platform (ISP) that allows the designers to assess the result of their design work assuming the point of view of the disabled. Within this last objective a set of custom developed Interaction Tools (ITs) have been developed. The novel ITs shall provide either sensorial stimulations altered according to the disability to be simulated with the aim of allowing designers to directly experience a specific disability, and therefore, get a deeper awareness and understanding of disabled user's needs, so as to help them in conceiving products with a better accessibility and usability [Hitc2001]. This set of ITs includes: Physical IT, Cognitive IT and Behavioral and Psychological IT. In this work we are focusing one subset of the Physical ITs that are Motor Functional Limitation IT that are concerned with simulation of motor impairments. In this paper we describe some preliminary developments for three types of ITs for different type of motor limitations simulation: Kinematics, Dynamics and Control Functional Limitation IT.

## 2. MOTOR FUNCTIONAL LIMITATION ITS

### 2.1 Scenario Description

In the foreseen scenario the designer will be interacting with a Virtual Reality system equipped with multiple retro projected walls and 3D stereo displays. The designer will be immersed in a virtual environment that reproduces the real scene of the product/application that she/he has previously designed. He will be able to

naturally interact with the environment thanks to wireless body tracking systems and through the purposely developed ITs. Before starting the immersive simulation she/he will be able to select a certain type of disabled user from a User Model database that is a database where the types of disabled user with the associated data are stored. The software platform will then load the parameters that are needed to be set for the initialization of the ITs according to the selected user. When the simulation starts the designer will be able to perform a virtual functional test of the designed product..

## 2.2 Motor Limitation Interaction Tools

Three kind of IT have been conceived for the simulation of three different type of motor disability.

### 2.2.1 Kinematic Functional Limitation IT

This tool implements a simulation of disabilities related to limited mobility impairment; the simulation is based on vibration feedback that occurs when the designer movements are not compatible with limitation of the simulated disabled user. According to value of the hands position retrieved by a body tracking system, whenever the arm/leg position exceeds the allowed range, a vibrotactile device generates a vibration directly on the joint of the body.

The vibrations are transmitted through a custom developed wireless module that makes use of eccentric mass vibration motors. In fig. 1-a one prototype of the vibrating module is shown. It consists of a bracelet equipped with four vibration element located around the wrist. The custom developed electronics is able to drive pulsing intermittent vibration with variable intensity (0-2N) and pulsing frequency (0-20Hz).

### 2.2.2 Dynamic Functional Limitation IT

This tool implements the simulation of disabilities related to functional limitations on the magnitude of the force needed to manipulate an object; the simulation is based on a specific tool, a force feedback steering wheel. Depending on the User Model data and on specific scenario data, extra forces will be applied to the steering wheel so as to let the designer perceive an artificial “weakness” simulating the functional disability. A customized steering wheel (see fig.2-b) able to exert strong torques up to 10Nm has been custom developed.

### 2.2.3 Control Functional Limitation IT

This tool implements simulation of disabilities consisting in functional limitations on the control of the movements; the simulation is based on a haptic device called GRAB [Aviz2003] which interferes with the designer movement introducing disturbing forces, on a selected bandwidth of frequency, simulating the functional limitation. The device is able to exert on the user wrist or finger an oscillating force in order to interfere with the natural movements simulating disabilities like pathological tremor.

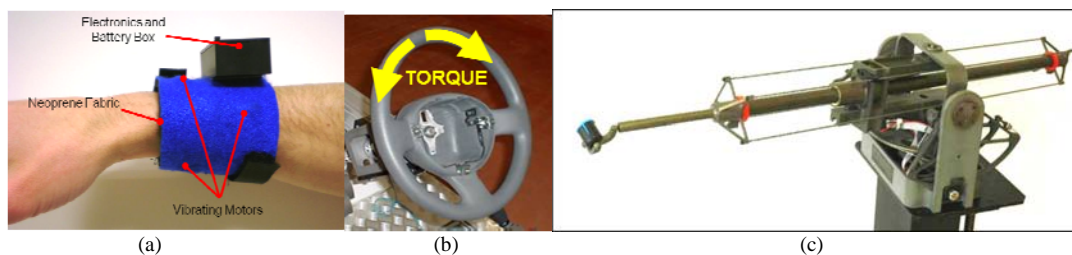


Figure 1. (a) First prototype of the vibrotactile bracelet equipped with 4 vibrating motors; (b) of a custom Force-Feedback steering wheel; (c) GRAB: Haptic user interface is able to deliver a force along any wanted orientation in the 3d space

### 3. CONCLUSION FUTURE WORKS

In this paper we introduce the three novel tools that have been developed within the Project VERITAS for disabilities simulation. The objective is to provide designer of products and service with direct experience of the disability of the end-user in order to get conscious of the disabled point of view. Three kind of novel/customized devices for specific type of motor disabilities have been developed and are currently being tested. In the very near future we are going to carry out a set of experiment for validation and parameter tuning the devices.

### ACKNOWLEDGEMENT

The work described here is funded by the EU Project N.247765 - VERITAS Integrated Project (IP) within the 7th Framework Programme, Theme FP7-ICT-2009.7.2, Accessible and Assistive ICT.

### REFERENCES

- Pheasant, S., 1996. *Bodyspace*, London: Taylor & Francis.
- Cardoso, Carlos; Clarkson, P. John. Simulation in user-centred design: helping designers to empathise with atypical users, *Journal of Engineering Design*. 29 Apr. 2011
- Hitchcock, D. R., Lockyer, S., Cook, S. and Quigley, C., 2001: Third age usability and safety -- an ergonomics contribution to design. In *International Journal of Human-Computer Studies*, 55 (4) pp. 635-643
- Avizzano, C.A.; Marcheschi, S.; Angerilli, M.; Fontana, M.; Bergamasco, M; *A Multi-Finger Haptic Interface for Visual Impaired*. Proceedings of the 12th IEEE Workshop Interactive Communication, October 31 - November 2, 2003, Millbrae, California US, ISBN: 0-7803-8136.



# **Doctoral Consortium**



# SENSORY FEEDBACK FOR MOTOR PERFORMANCE IN VIRTUAL ENVIRONMENTS ACROSS THE LIFESPAN

Patrick J. Grabowski and Andrea H. Mason  
*University of Wisconsin-Madison Department of Kinesiology  
2000 Observatory Drive, Madison, WI 53706*

## ABSTRACT

The results reported here are from a preliminary analysis of work in progress to investigate the use of visual feedback about one's own hand for motor control in virtual environments, and how that use changes as a function of age. Subjects from four age groups performed a reach to grasp task in a desktop virtual environment. Results indicate that older adults are not sensitive to visual feedback of their hand for reaching or grasping. Further, an augmentation effect was noted with regard to grasp aperture across all age groups. Research questions for future doctoral work are discussed.

## KEYWORDS

Motor control, virtual, reach, grasp, age

## 1. INTRODUCTION

Today, computer users include people from all walks of life, and all age groups. Interestingly, although computers have been commonplace in homes and work-environments for decades, the literature on interface design as it relates to age is only recent and limited in scope. In the late 1990's, interest in age-specific design increased, producing a reasonable body of knowledge on the design of standard desktop computer interface systems for various age groups. While the bulk of this information relates to cognitive performance (Hawthorn 2007), there is a modest body of scientific literature exploring motor control for human computer interaction (HCI) as a function of age (for example see Laursen, Jensen and Ratkevicius 2001).

Contrary to progress in age-specific design for standard computer systems, little is known about the age-related variance of motor output for HCI within three-dimensional (3D) virtual environments (VEs). In a series of experiments, McDonald et al. studied age-related differences in manipulation and locomotion. They concluded that, for younger females and older adults, performance deficits were likely influenced by limited video game and computer experience, as well as age-related changes in visual and motor systems (McDonald et al. 1999). Recent research has revealed age-related neurologic changes that lead to differences in the sensory and motor function of both elderly individuals and children when compared to young adults. Children and older individuals have fewer attentional resources and difficulty integrating multimodal sensory information for motor output (for example see Schut 1998). One important implication of these findings is that these individuals are more dependent on their visual systems for the control of movement (Adamo, Martin and Brown 2007). Because many virtual and augmented systems are being designed with the intention that the human hand will serve as the input device (Christou, Ritter and Jacob 2008), the provision of age-specific sensory information to users becomes an important issue. Therefore, the purpose of this doctoral research is to accumulate knowledge regarding the type of visual information required, when that information should be made available, and what form it should take for users of all age groups to effectively function in 3D VEs.

The specific research question for the current study relates to the utilization of visual feedback of one's own hand when performing a reach to grasp task in a VE. Previous research indicates that young adults are sensitive to the presence and timing of visual feedback about their hand (Mason and Bernardin 2008). Specifically, young adults show performance deficits in a reach to grasp task when vision of the hand is not present. These deficits are attenuated when even crude graphical feedback about the hand is provided. The timing of presentation of graphical feedback is also important, as information about hand location prior to



movement onset results in performance enhancement when compared to a condition where no information is provided. The current study will extend this prior research to gain knowledge on how the use of visual feedback changes as a function of age. Our hypothesis is that both children and older adults will be more sensitive to visual feedback of their own hand than younger and middle aged adults.

## 2. METHODS AND DESIGN

Data collection was conducted in the Wisconsin Collaborative Virtual Environment (WiscCVE – Figure 1A). An image of a target cube was displayed on a downward facing monitor. The image on the screen was reflected in a half-silvered mirror, appearing to the subject to be located in a workspace on the desktop surface (Figure 1B and C). Subjects wore CrystalEYES™ liquid crystal goggles to provide a stereoscopic view of the VE. Light-emitting diodes (LEDs) were fixed to the goggles, hand and target object (wooden cube 27 cm<sup>3</sup>, Figure 1D). These were tracked by a VisualEyez 3000 motion capture system (Phoenix Technologies, Inc) at a sampling rate of 120 Hz for scene rendering and data acquisition. This information was processed with a 10 ms lag to provide the subject with a real time, head-coupled view of the VE. The tracked position of LEDs on the physical target cube was used to generate the superimposed graphical image, which had identical dimensions to its physical counterpart. Traditional kinematic measures of human performance were used to evaluate the movements under the manipulated conditions. Overall movement time (MT) provides a general description of upper extremity movements, with longer times indicating greater task difficulty. For tasks involving grasping, peak grasp aperture (PA) is a standard measure. Larger apertures suggest a more complex task with greater attentional demand. Ultimately, complete analysis will include more complex kinematic measures (MacKenzie, C.L. and Iberall, T. 1994).

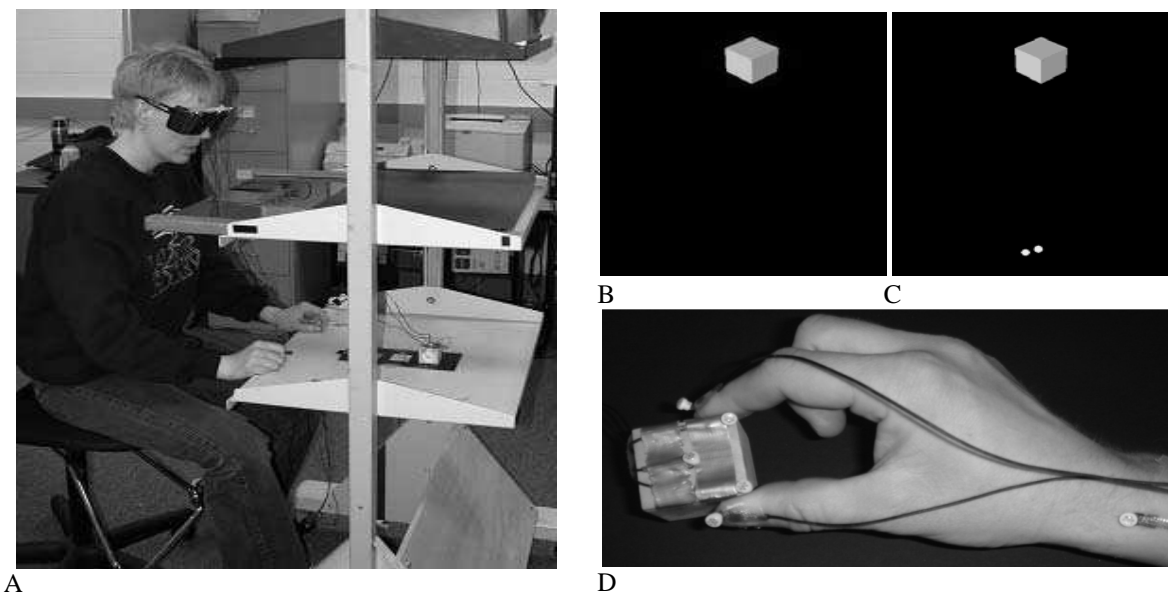


Figure 1. A) WiscCVE B) The visual scene without finger representation C) The visual scene with spheres present to represent fingertips D) The hand and target object instrumented with LEDs

The subject's task was to reach from a designated start position to grasp and lift a target cube. We manipulated target distance, visual feedback of the hand and age group membership. The target was placed at a 45-degree angle at either 7.5cm or 15cm from the start mark. At each location, the subject was provided with one of five visual feedback conditions of the hand. In the no vision (NV) condition, the subject was not given any graphical feedback about the position of the hand. In the full vision crude (FVC) condition, graphical feedback about hand position (10mm spheres at the fingertips) was provided throughout the entire reach-to-grasp movement. For the vision up to peak velocity (VPV) condition, graphical feedback about hand position was extinguished once peak velocity of the wrist was reached. In the vision until movement

onset (VMO) condition, graphical feedback of the hand was extinguished at the start of movement. For these conditions, subjects were prevented from seeing the real workspace below the mirror so that vision of the actual limb and surrounding environment was absent. For the final condition (full vision or FV), subjects were given full vision of the real hand as in a natural environment. Computer rendered graphical information about the target size and location was always available. Four age groups were utilized: children (7-12 years), young adults (18-30 years), middle age adults (40-50 years) and senior adults (60+ years). Forty-eight subjects, 12 in each age group participated. Subjects were seated in a height adjustable chair in front of the workspace. Subjects were instructed to reach and grasp the target at a comfortable pace. They were asked to lift the block vertically off the desktop once they had acquired it in a stable grasp. The experimenter then provided a verbal “done” cue and the subject set the block down and brought the hand back to the rest mark. The protocol was approved by the University of Wisconsin-Madison Social and Behavioral Sciences Institutional Review Board.

### 3. RESULTS

For each age group, data were analyzed using separate two-way (distance X visual feedback) repeated measures ANOVAs. For MT, there was a main effect of distance for all age groups (table 1). There was also a main effect of visual condition for all groups except seniors (Figure 2). For PA, there was a main effect of distance for all groups except children (table 1). There was a main effect of visual condition for young and middle age adults (Figure 3). All pairwise comparisons were performed with the Bonferonni method.

Table 1. Main effect of movement distance on MT and PA (mean±SE)

Age Group	Distance	Distance		F <sub>1,11</sub> p	Age Group	Distance	Distance		F <sub>1,11</sub> p
		Close	Far				Close	Far	
Children	MT	1046±54	1210±68	17.683, 0.001	Children	PA	70.9±3.1	70.2±2.1	0.276, 0.610
Young Adult	MT	766±71	913± 81	36.961, <0.001	Young Adult	PA	75.6±2.3	72.9±2.5	9.349, 0.011
Middle Adult	MT	877±84	1087±105	62.251, <0.001	Middle Adult	PA	71.3±2.3	74.1±2.3	22.995, 0.001
Senior Adult	MT	794±64	947± 88	13.703, 0.003	Senior Adult	PA	82.5±2.3	85.8±2.7	6.636, 0.026

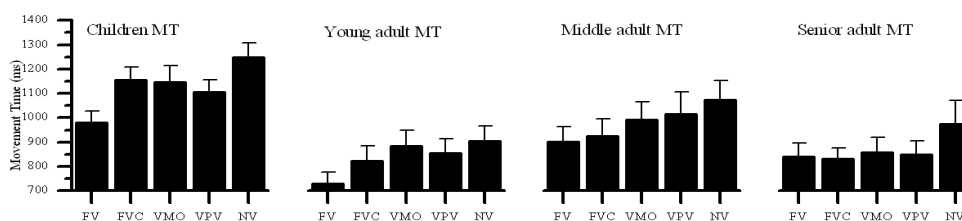


Figure 2. Effect of visual condition on movement time (mean and standard error): A) children  $F_{4,44}=10.088, p<0.001$  B) young adults  $F_{4,44}=7.241, p=0.001$  C) middle age adults  $F_{4,44}= 4.231, p=0.006$  D) senior adults  $F_{4,44}= 1.860, p=0.186$ . For children and young adults, having FV was significantly faster than NV ( $p<0.05$ ). In young adults, FV was also faster than VMO. There were no statistically significant differences between conditions in the middle and older age groups.

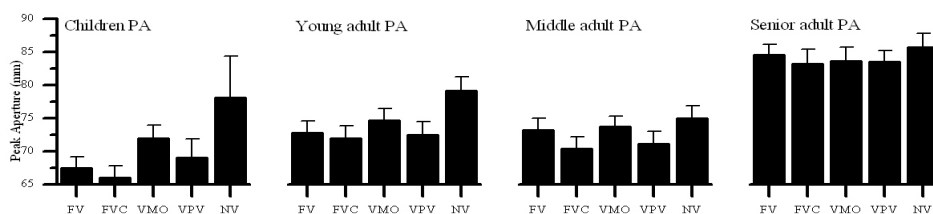


Figure 3. Effect of visual condition on peak aperture (mean and standard error): A) children  $F_{4,44}=3.061, p=0.102$  B) young adults  $F_{4,44}=10.011, p<0.001$  C) middle age adults  $F_{4,44}=4.144, p=0.022$  D) senior adults  $F_{4,44}=1.207, p= 0.320$ . For young and middle age adults, aperture in FVC was significantly smaller than in NV. For children and seniors FVC also had the smallest mean aperture, although this did not reach significance at the  $p<0.05$  level. Also of note is that in young adults, FV and VPV also had significantly smaller apertures than NV. Note the general lack of visual condition effect among seniors.

## 4. DISCUSSION

The most striking result is that older individuals are influenced very little by visual feedback about their hand. The hypothesis that these individuals will be more sensitive to visual information clearly does not pertain to visual feedback of self. The most plausible explanation for this finding is the decrease in attentional resources noted with increasing age (Schut 1998). Due to this limitation, older adults may not be able to visually attend to both their moving hand and a target to be grasped during the execution of a trial. While current position of the hand can be determined proprioceptively, it is well known that proprioceptive feedback declines with age (Adamo, Martin and Brown 2007). Thus, it is unlikely that proprioceptive feedback is providing them with the information necessary to complete the task. Further, studies of sensory integration indicate that older adults do not effectively combine visual and proprioceptive modalities for upper extremity tasks (Proteau, Charest, and Chaput 1994). In addition, functional multimodal sensory integration appears necessary for self-awareness (for review see Blanke 2009). Therefore, if self-feedback is of little consequence to older adults, what exactly are they looking at? In the current experiment, the only graphical images in the environment were the spheres representing the fingertips and the target cube. Further investigation into visual characteristics of the target is needed.

Formation of the grasp aperture also indicates two important findings. First, there is a very clear age difference in motor performance. Older adults consistently utilized a more conservative grasp formation, as evidenced by the wider aperture. Second, the full vision crude condition (finger representation present) had the smallest mean aperture in all groups (significant in two of four groups). The general suggestion from this result appears to be an augmentation effect. When the finger spheres were present throughout, they provided an advantage for movement execution. Specifically, a narrow aperture suggests that the movement may be more easily executed. The finger spheres in our environment may have provided a greater luminance contrast between the hand and target/background than provided by the other vision conditions, even the FV condition. The fact that performance was augmented by a visual representation with higher contrast suggests that further study on the effect of visual contrast of the effector for motor performance is warranted. In particular, the following research questions are suggested by the current results: Do older adults preferentially use visual information about the target object over information about self for motor output in VEs? Does the luminance contrast component of feedback about the hand serve as a constraint on movement execution? If so, is this a result of efficiency in the neural processing of visual feedback, a process known to vary with age? Can senior adults utilize visual information about self for motor performance if luminance contrast of the hand is maximized? Can age-related differences in performance in virtual and augmented environments be attenuated through use of high contrast visual feedback? Answers to these questions will lead to improvements in age-specific design of virtual and augmented environments.

## 5. CONCLUSION

The cross-sectional design of this study allows for confounding effects of generational differences in computer experience. However, we feel this is sufficiently limited by the use of a tangible user interface. One advantage of this interface is the use of natural movements without the need for abstraction to interact with the system. Since the experimental task in this study utilized the common act of reach to grasp, it is likely that computer experience played a minimal role in the outcome. This preliminary analysis indicates that for motor performance in a VE, senior adults do not effectively utilize visual feedback about their own hand for reaching and grasping. For grasping, an augmentation effect of graphic visual feedback was noted in all age groups. The information gained from this experiment is applicable for age-specific VE design, contributing to the fields of computer science, rehabilitation, geriatrics, and neuroscience.

## ACKNOWLEDGEMENTS

We would like to thank Brandon Bernardin, Drew Rutherford and Alexandra Skogen for support with data collection and analysis. This work was funded by the National Science Foundation grant No. 0916119.

## REFERENCES

- Adamo, D.E., Martin, B.J. and Brown, S.H. 2007. Age-related differences in upper limb proprioceptive acuity. *Perceptual and motor skills*, vol. 104, no. 3 Pt 2, pp. 1297-1309.
- Blanke, O. & Metzinger, T. 2009. Full-body illusions and minimal phenomenal selfhood. *Trends in cognitive sciences*, vol. 13, no. 1, pp. 7-13.
- Christou, G., Ritter, F.E. and Jacob, R. 2008. Modeling prehensile actions for the evaluation of Tangible User Interfaces. *Information Technology Interfaces, 2008. ITI 2008. 30th International Conference on IEEE*, Dubrovnik, Croatia , pp. 415-420.
- Hawthorn, D. 2007. Interface design and engagement with older people. *Behaviour and Information Technology*, vol. 26, no. 4, pp. 333-341.
- Laursen, B., Jensen, B.R. and Ratkevicius, A. 2001. Performance and muscle activity during computer mouse tasks in young and elderly adults. *European journal of applied physiology*, vol. 84, no. 4, pp. 329-336.
- MacKenzie, C.L. and Iberall, T. 1994. *The Grasping Hand.*, Elsevier Science, Amsterdam.
- Mason, A.H. and Bernardin, B.J. 2008. Vision for performance in virtual environments: The role of feedback timing. *Journal of Sport and Exercise Psychology*, 30, suppl S63.
- McDonald, D.P. et al, 1999. Performance differences between older and younger adults for a virtual environment locomotor task. *Automation technology and human performance: Current research and trends.* pp. 262-269.
- Proteau, L., Charest, I., and Chaput, S. 1994. Differential roles with aging of visual and proprioceptive afferent information for fine motor control. *Journal of Gerontology*. vol. 49, no. 3, pp. 100-107.
- Schut, L.J. 1998. Motor system changes in the aging brain: what is normal and what is not. *Geriatrics*, vol. 53 Suppl 1, pp. S16-9.

# NEURAL BASIS OF COMMUNICATION OF AUDITORY BCIs

Angela Riccio<sup>1</sup>, Luca Simione<sup>2</sup>, Donatella Mattia<sup>1</sup>, Febo Cincotti<sup>1</sup> and Marta Olivetti Belardinelli<sup>2</sup>

<sup>1</sup>*Neuroelectrical Imaging and BCI Laboratory, Fondazione Santa Lucia IRCCS, Rome, Italy*

*"Sapienza" Università di Roma*

<sup>2</sup>*ECoNA - Centro Interuniversitario di Ricerca sull'Elaborazione Cognitiva nei Sistemi Naturali ed Artificiali.*

## ABSTRACT

The aim of this project is a systematic and critical analysis of the studies exploring different interaction modalities in Brain Computer Interfaces (BCIs) for communication.

BCIs measure signals related to specific brain activity and translate them into device control signals. This technology can be used to provide users with locked-in syndrome (e.g. late stage Amyotrophic Lateral Sclerosis, ALS) with an assistive device that does not rely on muscular contraction. Several studies about BCI explored mental tasks and paradigms using visual modality. Considering that in ALS patients the oculomotor control can deteriorate and also the potential users with other diseases could have impaired visual function, in the last years, tactile and auditory modality have been investigated to find alternative BCI independent of vision.

The aim of this review is to survey the cognitive and neural basis of the tasks explored to control the auditory BCI systems to better understand their limits and potentials applications, toward a real user centered approach helping the scientific community to move the BCIs from the laboratory to user's houses.

## KEYWORDS

Brain computer interface, auditory, communication, P300, ERPs, neurophysiology.

## 1. INTRODUCTION

Brain Computer Interface (BCI) is a technology that utilizes neurophysiological signals from the brain to control external devices (Wolpaw et al, 2002; Birbaumer, 2006).

Many BCI systems utilize neurophysiological signals recorded by electroencephalography (EEG; see Birbaumer 2006 for a review) and most of the studies about the EEG-based BCI are focused on the vision as interaction modality. Often they are implemented in an oddball paradigm, wherein rare target are intersected with frequent non target events. This current BCI applications are based on the non-invasively measurable P300 evoked potential as input signal. The P300 event related potential is a positive deflection in the human EEG, appearing approximately 300 ms after the presentation of rare or surprising, task relevant stimuli (Sutton et al., 1965).

## 2. INTERACTING WITH BCI SYSTEMS

### 2.1 Visual-based BCIs

Farwell and Donchin (1988) introduced the first visual P300-based BCI paradigm (P300 speller). The P300 speller consists of a 6x6 symbol matrix wherein symbols are arranged within rows and columns. Throughout the course of a trial, rows and columns are intensified one after the other in a random order. The user is asked to focus attention on the one of the 36 cells with the desired letter and mentally count the number of time it flashes. The computer identifies the attended item as the intersection of the row and column that elicited the largest P300. This stimulus is the "rare event" in an oddball paradigm (Fabiani et al., 1987) and elicits the p300 of ERPs.

Afterwards, many studies were conducted on the P300 speller (Farwell & Donchin, 1988; Nijboer et al., 2008; Kleih et al., 2010) and also similar visual p300-based device have been developed and investigated with able and not able-bodied subjects (Sellers and Donchin, 2006; Piccione et al., 2006; Hoffman et al., 2008). However, most of the tested devices rely to some extent on the ability of the subject to control his own eyes. Two recent studies pointed out how the ability to control such a visual P300-based BCI depends on gaze direction (Brunner et al. 2010; Treder and Blankertz, 2010). Indeed the performances on the classical p300 visual speller deteriorates if one switches from overt to covert attention (Brunner et al., 2010; Treder & Blankertz, 2010).

## 2.2 Different Modalities for Interaction with BCI

The primary motivation of BCI research is to reestablish communication and facilitation of daily life activities for people with severe motor disability due to different neurological disorders such as ALS, acquired brain injury, cerebral palsy, spinal cord injuries, muscular dystrophies or chronic peripheral neuropathies. The BCI literature showed that potential end users in the locked-in state (LIS) are able to successfully control a visual-based BCI (Kubler et al., 2001, 2007; Neuper et al., 2003) However the potential end users in complete locked-in (CLI) failed to achieve BCI control. Considering that in ALS patients in a progressed stages of the disease the oculomotor control can deteriorate and also oculomotor and visual disease can affect the Acquired Brain Injury potential users, developing new BCI paradigm that are not dependent on eye gaze is one of the challenging issues in modern BCI research (Nijboer et al., 2009).

In this regard Treder and Blankertz (2010) presented an ERP device which can be used in a covert attention condition, in which subject attention is moved to a target without gazing to it. Around one circle in the middle of the screen, six hexagons were arranged. Each of the hexagons contained 5 letters of the alphabet and was randomly highlighted. When chosen, this hexagon expanded the letters contained in it into the other six hexagons in one empty hexagon. In covert attention condition performances up to 60% were found using this stimulation modality. Also other studies proposed some visual interfaces with unusual organization of the stimuli and feedback, to make the BCI visual devices independent from the ability to gaze (Liu et al., 2010, Aloise et al., 2010, Acqualogna et al., 2010).

On the other hand different interaction modalities in BCI applications, such as the auditory (Hadler et al., 2010, Klobassa et al., 2010, Schreuder et al., 2010) and the tactile ones (Muller-Putz et al, 2006; Cincotti et al., 2007; Brouwer and van Erp, 2008) have been explored. Here we will focus on the auditory P300 based BCI paradigm.

## 2.3 Auditory BCIs

To control an auditory P300-based BCI system participants are asked to concentrate on a target stimulus among non-target stimuli, such as tones, words, or numbers, acoustically presented. One of the first studies exploring the auditory modality in the BCI field (Sellers and Donchin, 2006) used the words “Yes”, “No”, “Stop”, “Pass” as possible targets. The authors compared the auditory, visual and auditory + visual stimulation mode along 10 experimental sessions. The three different stimulation mode were investigated with three healthy subjects and three ALS patients. On average the subjects achieved an offline classification performance accuracy of 65% in the auditory modality. Furdea et al. (2009) also used words acoustically presented as BCI stimuli. They assigned numbers to identify rows and columns of a 5x5 speller matrix and such numbers were presented for the selection. In order to select a letter the users were asked to attend to the two numbers representing the coordinates of the target character in the matrix. The authors compared the auditory and visual modality within 15 healthy users who performed with an online accuracy of 65% for the auditory and the 94.62% for the visual. Other studies used sounds as stimuli instead of words. Kloblassa et al. (2009), following the same approach of Furdea et al. (2009), coded a 6x6 P300 speller with different environmental sounds instead of numbers and run a test with 10 healthy users along 11 sessions. Five subjects tested directly the auditory modality speller, whereas the other five tested the speller initially with both auditory and visual modalities, then the visual stimuli were gradually withdrawn until only the auditory stimuli remained. No statistically significant differences in the online performances were found both in the first and in the last session between the two groups. The mean of the accuracy for the final session with auditory stimuli was 59.38 % (SD 22.10) for the first group and 68.78 % (SD 13.98) for the second group. A

comparison of the two modes (auditory only vs. auditory plus visual) also showed higher accuracies in concurrent auditory and visual stimulus presentation than in auditory stimulus presentation. Considering that 70% is the lower acceptable limit for spelling with an assistant device for communication (Kubler et al., 2006) the average accuracy obtained in the former studies was still too low.

Other studies with auditory BCI paradigms used tones as stimuli instead of words. Schreuder et al. (2010) proposed a multi-class auditory paradigm that uses spatially distributed cues. The subjects were surrounded by eight speakers at ear height, spaced with 45° angle between them, at approximately one meter distance from the subject's ears. By presenting to ten healthy subject an oddball task with the spatial location of the stimuli being a discriminative cue in an offline study, they reached over than 90% of selection score. All subjects were able to perform a selection score of over than 70% in the conditions with spatial cues and all but one subject reached a selection score higher than 90%.

Despite of the good results obtained by Schreuder et al. (2010), Halder et al. (2010) underlined that some users could be incapable of using a multiclass auditory P300, e.g. looked-in patients who suffer from reduced attention span (Birbaumer, 2006). They thus proposed an auditory BCI based on a three stimuli paradigm for binary selection. The authors enrolled 20 healthy subjects and presented them with three different tasks of three stimuli oddball that varied for pitch, loudness and direction of the target tones. They founded that the pitch was of superior efficacy for most participants, but some participants performed better in the other tasks. The mean offline classification accuracy of the best task of each subject was 78.54 %. On the basis of this result, they concluded that a screening session to determine the optimal task for each user is recommended, to provide the better solution for each user.

## 2.4 Auditory P300 and its Neural Correlates in Human Brain

Differently from the visual P300-based BCIs, few studies using auditory P300-based BCIs were conducted on potential users (Seller and Donchin, 2006; Kubler et al., 2009). Kubler and colleagues (2009) tested the same 5x5 auditory matrix presented in Furdea et al., (2009) with four ALS patients. The online selection accuracies for the four patients were respectively of 25%, 0%, 0% and 23.53%. The offline classification accuracies were of 58.33%, 25%, 25% and 41.17%. The performances registered with this auditory BCI appear very low, especially if compared with previously recorded data collected with the same patients using a visual P300-based BCI (Nijboer et al., 2008). Seller and Donchin (2006) tested a four-choice auditory based BCI on three ALS patients. The offline classification accuracies of the three patients, averaged across all 10 experimental sessions, were of 65.4%, 73.2% and 59.1%. Thus, the performances of two of these patients were not high enough to allow meaningful communication (selection accuracy of at least 70% , Kubler et al. 2001),

Clinical evidence of cognitive impairment can be detected in up of 50% percent of patients with ALS through direct neuropsychological testing although frontotemporal dementia (FTD) occurs in a limited percentage of these patients (Woolley and Katz, 2008). Several studies showed that the primary deficit in ALS occur in the domain of attention and cognitive flexibility (Massman et al., 1986; Frank et al., 1997). These data pointed out the importance of taking into account the patient difficulties in controlling a BCI, due to brain injuries or cognitive impairment, in development and evaluation processes of an auditory-based BCI system.

The understanding of the difficulties of the users in controlling the auditory P300-based BCI would be greatly increased if the auditory P300 component could be associated with specific neural structures or circuits in human brain (for a review O' Donnell et al., 1999). Different approaches have been used to study neural structures associated with the scalp-recorded auditory P300, e.g. by using depth recordings and cases of brain lesions in order to identify regions that might contribute to this component. These studies founded that the temporal lobe and the inferior parietal lobe structures are involved in the generation of the auditory P300 (O'Donnell et al., 1993). Functional MRI studies using auditory oddball paradigms have also implicated the temporal-parietal cortex. Turetsky, Raz, and Alsop (1998) reported activation in the superior temporal gyrus, and Menon, Ford, Lim, Glover, & Pfefferbaum (1997) reported activation in the supramarginal gyrus of the parietal lobe.

### 3. CONCLUSION

Generally, studies on BCI technology are focused on the accuracy of the system, how fast it can be trained, how long it takes to prepare the subjects (Thulasidas & Guan, 2005), the effects of different electrode configurations ( Krusiensky et al., 2008), machine learning algorithms (Hoffman et al. 2006), stimulus onset asynchrony (Allison & Pineda, 2006) and the modality of stimulus presentation (Allison & Pineda, 2006; Townsend et al. 2010) on classification accuracy.

Here we propose a review which will move the focus from these technical aspects to the potential users' cognitive and sensorial abilities implied in the BCI tasks. Indeed, we are interested in deepening the neural basis of such cognitive abilities, to better understand how to improve the reviewed auditory-based BCI systems. To consider neurophysiological substrates and cognitive stimuli processing in different sensorial modalities is an important step for the BCI research. It would help to understand more about the limits of the developed technology and the reasons why some of the users are not able to control them and consequently to adapt the stimulation paradigm.

### ACKNOWLEDGEMENT

This work is supported by the European ICT Programme Project FP7- 224631 (TOBI). The text reflects solely the views of its authors. The European Commission is not liable for any use that may be made of the information contained therein.

### REFERENCES

- Acqualagna L. et al, 2010. A novel brain-computer interface based on the rapid serial visual presentation paradigm. *Conf Proc IEEE Eng Med Biol Soc.* no.1, pp 2686-9
- Allison B.Z. & Pineda J.A. 2006, Effects of SOA and flash pattern manipulations on ERPs, performance, and preference: implications for a BCI system. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology* 59 no. 2 pp. 127-140.
- Birbaumer N. 2010, Brain-computer-interface research: coming of age. *Clin Neurophysiol* 117 no 3 pp. 479-83
- Birbaumer N. et al, 2008 Brain-computer interface in paralysis. *Curr Opin Neurol.*, 21, no. 6, pp. 634-638.
- Brouwer A.M. & van Erp J. 2008 A tactile P300 BCI and the optimal number of factors: effects of target probability and discriminability. In: *Proceedings of the 4th International Brain-Computer Interface Workshop and Training Course 2008.* TU-Graz. pp. 280-285
- Brunner P. et al., 2010. Does the 'P300' speller depend on eye gaze? *Journal of Neural Engineering*, no. 7(5), 056013.
- Cincotti F. et al. 2007 Vibrotactile Feedback for Brain-Computer Interface Operation. *Comput Intell Neurosci* Vol. 2000, 12 pages
- Farwell L.A. and Donchin E. 1988 Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalogr ClinNeurophysiol* no.70, pp 510-523.
- Furdea A. et al. 2009 An auditory oddball (P300) spelling system for brain-computer interfaces. *Psychophysiology*, 46 no 3, pp. 617-25.
- Frank B. et al. 1997, Relation of neuropsychological and magnetic resonance findings in amyotrophic lateral sclerosis: evidence for subgroups. *Clinical Neurology and Neurosurgery*, 99, no. 2, pp. 79-86.
- Halder S. et al. 2010. An auditory oddball brain-computer interface for binary choices. *Clinical Neurophysiology: Official Journal Of The International Federation Of Clinical Neurophysiology*, 121, no 4, pp 516-523.
- Hoffmann U. et al. 2008 An efficient P300-based brain-computer interface for disabled subjects. *Journal of Neuroscience Methods*, 167, no. 1, pp. 115-125.
- Klobassa D.S. et al. 2009 Toward a high-throughput auditory P300-based brain-computer interface. *Clin Neurophysiol*, no.120, pp 1252-1261.
- Kubler A. et al. 2001 Brain-computer communication: self-regulation of slow cortical potentials for verbal communication. *Arch. Phys. Med. Rehabil.* no. 82, pp. 1533-1539
- Kubler A. et al. 2009, A brain-computer interface controlled auditory event-related potential (p300) spelling system for locked-in patients. *Ann N Y Acad Sci* 2009 no. 1157 pp. 90-100.



- Liu Y. et al. 2010. Gaze independent brain-computer speller with covert visual search tasks. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21163695>
- Menon V. et al. 1997 Combined event-related and EEG evidence for temporal-parietal cortex activation during target detection. *Neuroreport*, no. 8, pp.3029–3037.
- Muller-Putz G.R., et al. 2006 Steady-state somatosensory evoked potentials: suitable brain signals for brain-computer interfaces? *IEEE Trans Neural Syst Rehabil Eng*, no.14, pp 30–37.
- Nijboer F. et al. 2008 A P300-based brain-computer interface for people with amyotrophic lateral sclerosis. *Clinical Neurophysiology Official Journal of the International Federation of Clinical Neurophysiology*, 119 no. 8, pp. 1909-1916.
- O'Donnell B.F. et al. 1999 Identification of neural circuits underlying P300 abnormalities in schizophrenia. *Psychophysiology* no. 36 pp. 388±398.
- Piccione F. et al. 2006 P300-based brain computer interface: reliability and performance in healthy and paralysed participants. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 117, no 3, pp 531-537.
- Schreuder M. et al. 2010 A new auditory multi-class brain-computer interface paradigm: spatial hearing as an informative cue. *PLoS One*. Apr no.1;5(4) e9813.
- Sellers E & Donchin E. 2006 A P300-based brain-computer interface: initial tests by ALS patients. *Clin Neurophysiol*, 117 no. 3, pp. 538–48.
- Sutton S. et al. 1965 Evoked-potential correlates of stimulus uncertainty. *Science (New York, N.Y.)*, no. 150(700), pp. 1187-1188.
- Thulasidas M. & Guan C. 2005 Optimization of BCI Speller Based on P300 Potential. *Conference Proceeding IEEE Engineering in Medicine and Biology Society*. no. 5, pp. 5396-9.
- Treder M.S. & Blankertz B., 2010 (C)overt attention and visual speller design in an ERP-based brain-computer interface. *Behavioral and Brain Functions: BBF*, no. 6, pp 28.
- Townsend G. et al. 2010 A novel P300-based brain-computer interface stimulus presentation paradigm: moving beyond rows and columns. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 121 no. 7, pp. 1109-1120.
- Turesky B.I. et al. 1998 Functional magnetic resonance imaging of the sources of the auditory P300. *Paper presented at the 12th International Congress of Evoked Potentials*, Boston. 1998 Jul.
- Wolpaw J.R. et al., 2002. Brain-computer interfaces for communication and control. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 113 no. 6, pp. 767-791
- Woolley S.C. & Katz J.S. 2008, Cognitive and behavioral impairment in amyotrophic lateral sclerosis. *Physical Medicine and Rehabilitation Clinics of North America*, 19, no. 3, pp. 607-617.

# AUTHOR INDEX

Aguiar, Y. ....	105	Delaplace, S. ....	185
Ahn, Y. ....	327	Duarte, C. ....	69
Al-ani, T. ....	185	Dubinsky, Y. ....	442
Alhadreti, O. ....	3	Elbaz, P. ....	374
Alshamari, M. ....	3	Ertl, T. ....	113, 153
Aoki, H. ....	389	Evreinov, G. ....	209
Aoyama, H. ....	241	Evreinova, T. ....	209
Aubert, D. ....	438	Feng, W. ....	480
Back, J. ....	499	Flood, D. ....	217, 434
Baranauskas, M. ....	233	Fontana, M. ....	509
Becker, M. ....	355	Fontecha, J. ....	337
Belardinelli, M. ....	520	Frieß, M. ....	249
Betke, M. ....	177	Fujita, H. ....	471, 474
Bitarello, B. ....	384	Fujita, T. ....	489
Blascheck, T. ....	153	Fukazawa, Y. ....	121
Blum, R. ....	365	Galy, E. ....	105
Bold, D. ....	153	Gapenne, O. ....	438
Bomsdorf, B. ....	265	Garg, A. ....	417
Bordegoni, M. ....	201	Gelderblom, H. ....	10
Bouchet, F. ....	61	Gennari, R. ....	37
Bravo, J. ....	337	Germanakos, P. ....	283
Brayshaw, M. ....	193	Gesso, I. ....	53
Braz, A. ....	384	Gordon, M. ....	360
Burmeister, O. ....	409	Grabowski, P. ....	515
Burmester, M. ....	137	Granić, A. ....	495
Cabitza, F. ....	53	Grant, R. ....	225
Caelen, J. ....	355	Gu, B. ....	499
Capuano, D. ....	429	Guevara, J. ....	455
Carlini, M. ....	37	Gurari, D. ....	177
Carrozzino, M. ....	509	Gusmini, M. ....	417
Catarci, T. ....	442	Haag, F. ....	113
Chang, W. ....	447	Hanasaka, Y. ....	486
Chao, C. ....	480	Harel, D. ....	360
Chen, H. ....	483	Harrison, R. ....	217, 434
Chen, Y. ....	477	Harviainen, T. ....	129
Chien, S. ....	29	Hasan, L. ....	45
Chimbo, B. ....	10	Hasegawa-Ohira, M. ....	486
Cincotti, F. ....	520	Held, T. ....	257
Clarke, M. ....	365	Hervás, R. ....	337
Corna, S. ....	53	Hong, J. ....	327
Cosi, P. ....	369	Hong, S. ....	503
Covarrubias, M. ....	201	Hornung, H. ....	233
Cugini, U. ....	201	Huang, K. ....	447
De Monte, ....	429	Humayoun, S. ....	442
DeAngelis, D. ....	225	Ihamäki, P. ....	459

Ilmberger, W. ....	257	Mercantini, J. ....	105
Ishiguro, T. ....	486	Migge, B. ....	93
Işık, N. ....	86	Miranda, L. ....	233
Islam, M. ....	19	Miyoshi, T. ....	389, 399
Israel, A. ....	442	Monacelli, E. ....	185
Iwata, H. ....	121	Mora, S. ....	423
Jang, T. ....	499	Mornar, J. ....	495
Jemni, M. ....	185	Mourlas, C. ....	283
Juha, L. ....	161	Muhler, D. ....	153
Kalwar, S. ....	321	Muramatsu, N. ....	394
Kato, M. ....	492	Nakagawa, M. ....	399
Kawase, Y. ....	506	Nakajima, S. ....	474
Kieffer, S. ....	404	Nakashima, Y. ....	471, 474, 489
Kihara, T. ....	379	Nakatani, Y. ....	291, 379, 506
Kim, B. ....	177	Nakayasu, H. ....	389, 399
Kim, K. ....	503	Nakazawa, J. ....	316
Kleinhans, M. ....	249	Nam, K. ....	503
Knoch, S. ....	345	Nazarov, E. ....	442
Koskinen, H. ....	129	Neca, J. ....	69
Kovar, S. ....	341	Nganji, J. ....	193
Kröner, A. ....	345	Nomura, S. ....	486
Kunz, A. ....	93	O'Donnell, H. ....	177
Kuo, F. ....	299	Ogawa, H. ....	486
Laarni, J. ....	129	Ohshima, K. ....	241
Lai, C. ....	477	Önal, K. ....	86
Le, N. ....	316	Patterson, P. ....	389, 399, 413
Lee, H. ....	499	Paulheim, A. ....	265
Lekkas, Z. ....	283	Pellegrin, A. ....	355
Leone, G. ....	369	Perry, D. ....	225
Lin, C. ....	447	Pertti, S. ....	161
Lin, C. ....	447, 483	Pohl, M. ....	341
Liu, X. ....	438	Pontual, G. ....	311
Lohmann, A. ....	384	Potier, L. ....	438
Luu, D. ....	225	Raisamo, R. ....	209
Luzayisu, E. ....	273	Raschke, M. ....	153
Macq, B. ....	404	Rebekah, R. ....	161
Magin, D. ....	351	Reiszel, F. ....	384
Mahdavi, A. ....	29	Riccio, A. ....	520
Martin, C. ....	217, 434	Roccaforte, M. ....	429
Mascio, T. ....	37	Rocha, L. ....	384
Mason, A. ....	515	Rodriguez, N. ....	146
Masri, F. ....	423	Ryall, K. ....	225
Mast, M. ....	137	Sakamoto, M. ....	394
Masuda, M. ....	291	Samaras, G. ....	283
Matsuda, N. ....	169	Sansonnet, J. ....	61
Mattia, D. ....	520	Santoni, C. ....	105
Mayhew, P. ....	3	Sato, T. ....	492
McDaid, K. ....	217	Sawa, K. ....	471

Sbattella, L. ....	77
Schlegel, T. ....	153
Schmidt, T. ....	93
Schrepp, M. ....	257
Sendai, K. ....	474
Shimodaira, Y. ....	492
Shirogane, J. ....	121
Simione, L. ....	520
Sinzinger, E. ....	413
Steffen, D. ....	351
Sun, T. ....	480
Sun, T. ....	483
Swine, C. ....	404
Takadama, K. ....	394
Takamatsu, M. ....	471, 474, 489
Takashio, K. ....	316
Takeuchi, H. ....	169
Taras, C. ....	113
Tedesco, R. ....	77
Tokuda, H. ....	316
Tomasuolo, E. ....	429
Trad, D. ....	185
Trivilini, A. ....	77
Tseng, F. ....	480
Tsianos, N. ....	283
Tsikerdekis, M. ....	463
Turner, W. ....	61
Umemuro, H. ....	455
Vanderdonckt, J. ....	273
Velez, E. ....	146
Vieira, M. ....	105
Villiers, R. ....	10
Wu, M. ....	299
Wu, W. ....	477
Wu, Y. ....	477
Ximenes, V. ....	311
Yaremko, N. ....	332
Yashita, Y. ....	121
Zach, R. ....	29