22 Schriften aus der Fakultät Wirtschaftsinformatik und Angewandte Informatik der Otto-Friedrich-Universität Bamberg

INTERACT 2015 Adjunct Proceedings

15th IFIP TC.13 International Conference on Human-Computer Interaction 14 – 18 September 2015, Bamberg, Germany

Christoph Beckmann, Tom Gross (eds.)

Connection.Tradition.Innovation
INTERACT 2015

University of Bamberg Press

22 Schriften aus der Fakultät Wirtschaftsinformatik und Angewandte Informatik der Otto-Friedrich-Universität Bamberg

Contributions of the Faculty Information Systems and Applied Computer Sciences of the Otto-Friedrich-University Bamberg

Schriften aus der Fakultät Wirtschaftsinformatik und Angewandte Informatik der Otto-Friedrich-Universität Bamberg

Contributions of the Faculty Information Systems and Applied Computer Sciences of the Otto-Friedrich-University Bamberg

Volume 22

University of Bamberg Press
2015

INTERACT 2015 Adjunct Proceedings

15th IFIP TC.13 International Conference
on Human-Computer Interaction
14 – 18 September 2015, Bamberg, Germany

Christoph Beckmann, Tom Gross (eds.)

University
of Bamberg
Press
2015

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliographie; detailed bibliographic data are available on the Internet at http://dnb.d-nb.de/.

CR-Code to the Open Access Publication

This publication is available OPEN ACCESS on the institutional repository OPUS (http://www.opus-bayern.de/uni-bamberg/) of the University of Bamberg. For the print copies all rights are reserved according to the Copyright law of Germany.

Printing and Binding: Digital Print Group, Nürnberg
Cover-Design: University of Bamberg Press, Anna Hitthaler

© University of Bamberg Press Bamberg 2015
http://www.uni-bamberg.de/ubp/

ISSN: 1867-7401
ISBN: 978-3-86309-352-5 (Print)
eISBN: 978-3-86309-353-2 (Online)
URN: urn:nbn:de:bvb:473-opus4-256444

Preface

This volume presents the Adjunct Proceedings of the 15th IFIP TC.13 International Conference on Human-Computer Interaction – INTERACT 2015 held from 14 to 18 September 2015 in Bamberg, Germany. It contains the position papers for the students of the Doctoral Consortium as well as the position papers of the participants of the various workshops.

The Proceedings of the 15th IFIP TC.13 International Conference on Human-Computer Interaction – INTERACT 2015 published by Springer-Verlag contain the full research papers, short research papers, interactive posters and demos as well as tutorials.

INTERACT is among the world's top conferences in Human-Computer Interaction. Starting with the first INTERACT conference in 1990, this conference series has been organised under the aegis of the Technical Committee 13 on Human-Computer Interaction of the UNESCO International Federation for Information Processing (IFIP). This committee aims at developing the science and technology of the interaction between humans and computing devices.

The theme of INTERACT 2015 in Bamberg was "Connection.Tradition.Innovation". The city of Bamberg is proud of its more than 1000-year-old centre. It has more than 2,400 historically listed buildings and became a UNESCO World Cultural Heritage Site in 1993. With 70000 inhabitants, Bamberg is a small town in the heart of Europe. The University of Bamberg is an illustrative example of combining the past and the present—it is one of the oldest universities of Germany and was originally founded in 1647; after being closed since the 19th century, it became one of the newest universities of Germany to open again in 1979 and is, therefore, an ideal place to gather under the theme.

We would like to thank all authors of the papers of these adjunct proceedings as well as the co-chairs of these events and their committees and reviewers. We also thank Barbara Ziegler from the University of Bamberg Press for her great support during the production of this volume.

Christoph Beckmann, Tom Gross

Volume Editors

Christoph Beckmann
University of Bamberg
An der Weberei 5, 96047 Bamberg, Germany
christoph.beckmann@uni-bamberg.de

Tom Gross
University of Bamberg
An der Weberei 5, 96047 Bamberg, Germany
tom.gross@uni-bamberg.de

IFIP TC.13

Established in 1989, the International Federation for Information Processing Technical Committee on Human–Computer Interaction (IFIP TC.13) is an international committee of 37 member national societies and 9 Working Groups, representing specialists in human factors, ergonomics, cognitive science, computer science, design and related disciplines. INTERACT is its flagship conference, staged biennially in different countries in the world. From 2017 the conference series will develop to become an annual conference.

IFIP TC.13 aims to develop the science and technology of human–computer interaction (HCI) by encouraging empirical research, promoting the use of knowledge and methods from the human sciences in design and evaluation of computer systems; promoting better understanding of the relation between formal design methods and system usability and acceptability; developing guidelines, models and methods by which designers may provide better human-oriented computer systems; and, cooperating with other groups, inside and outside IFIP, to promote user-orientation and humanization in system design. Thus, TC.13 seeks to improve interactions between people and computers, encourage the growth of HCI research and disseminate these benefits worldwide.

The main orientation is toward users, especially the non-computer professional users, and how to improve human–computer relations. Areas of study include: the problems people have with computers; the impact on people in individual and organizational contexts; the determinants of utility, usability and acceptability; the appropriate allocation of tasks between computers and users; modelling the user to aid better system design; and harmonizing the computer to user characteristics and needs.

While the scope is thus set wide, with a tendency toward general principles rather than particular systems, it is recognized that progress will only be achieved through both general studies to advance theoretical understanding and specific studies on practical issues (e.g., interface design standards, software system consistency, documentation, appropriateness of alternative communication media, human factors guidelines for dialogue design, the problems of integrating multimedia systems to match system needs and organizational practices, etc.).

In 1999, TC.13 initiated a special IFIP Award, the Brian Shackel Award, for the most outstanding contribution in the form of a refereed paper submitted to and delivered at each INTERACT. The award draws attention to the need for a comprehensive human-centered approach in the design and use of

information technology in which the human and social implications have been taken into account. 2007 IFIP TC 13 also launched an Accessibility award to recognize an outstanding contribution with international impact in the field of accessibility for disabled users in human-computer interaction. 2013 IFIP TC 13 launched the Interaction Design for International Development (IDID) Award recognizes the most outstanding contribution to the application of interactive systems for social and economic development of people in developing countries. Since the process to decide the award takes place after papers are submitted for publication, the awards are not identified in the proceedings.

IFIP TC 13 also recognizes pioneers in the area of Human-Computer Interaction. An IFIP TC 13 Pioneer is one who, through active participation in IFIP Technical Committees or related IFIP groups, has made outstanding contributions to the educational, theoretical, technical, commercial or professional aspects of analysis, design, construction, evaluation and use of interactive systems. IFIP TC 13 pioneers are appointed annually and awards are handed over at the INTERACT conference.

IFIP TC.13 stimulates working events and activities through its Working Groups (WGs). WGs consist of HCI experts from many countries, who seek to expand knowledge and find solutions to HCI issues and concerns within their domains, as outlined below.

WG13.1 (Education in HCI and HCI Curricula) aims to improve HCI education at all levels of higher education, coordinate and unite efforts to develop HCI curricula and promote HCI teaching.

WG13.2 (Methodology for User-Centered System Design) aims to foster research, dissemination of information and good practice in the methodical application of HCI to software engineering.

WG13.3 (HCI and Disability) aims to make HCI designers aware of the needs of people with disabilities and encourage development of information systems and tools permitting adaptation of interfaces to specific users.

WG13.4 (also WG2.7) (User Interface Engineering) investigates the nature, concepts and construction of user interfaces for software systems, using a framework for reasoning about interactive systems and an engineering model for developing user interfaces.

WG 13.5 (Resilience, Reliability, Safety and Human Error in System Development) seeks a frame- work for studying human factors relating to systems failure, develops leading edge techniques in hazard analysis and safety engineering of computer-based systems, and guides international accreditation activities for safety-critical systems.

WG13.6 (Human-Work Interaction Design) aims at establishing relationships between extensive empirical work-domain studies and HCI design. It will promote the use of knowledge, concepts, methods and techniques that enable user studies to procure a better apprehension of the complex interplay between individual, social and organizational contexts and thereby a better understanding of how and why people work in the ways that they do.

WG13.7 (Human–Computer Interaction and Visualization) aims to establish a study and research program that will combine both scientific work and practical applications in the fields of Human–Computer Interaction and Visualization. It will integrate several additional aspects of further research areas, such as scientific visualization, data mining, information design, computer graphics, cognition sciences, perception theory, or psychology, into this approach.

WG13.8 (Interaction Design and International Development) are currently working to reformulate their aims and scope.

WG13.9 (Interaction Design and Children) aims to support practitioners, regulators and researchers to develop the study of interaction design and children across international contexts.

New Working Groups are formed as areas of significance to HCI arise. Further information is available at the IFIP TC.13 website: http://ifip-tc13.org/

IFIP TC.13 Members

Officers

Chairperson
Jan Gulliksen

Vice-chair
Philippe Palanque

Vice-Chair for WG and SIG
Simone D. J. Barbosa

Treasurer
Anirudha Joshi

Secretary
Marco Winckler

Webmaster
Helen Petrie

Country Representatives

Australia
Henry B.L. Duh
Australian Computer Society

Austria
Geraldine Fitzpatrick
Austrian Computer Society

Belgium
Monique Noirhomme-Fraiture
Fédération des Associations Informatiques de Belgique

Brazil
Raquel Oliveira Prates
Brazilian Computer Society (SBC)

Bulgaria
Kamelia Stefanova
Bulgarian Academy of Sciences

Canada
Heather O'Brien
Canadian Information Processing Society

Chile
Jaime Sánchez
Chilean Society of Computer Science

China
Chengqing Zong
Chinese Institute of Electronics

Croatia
Andrina Granic
Croatian Information Technology Association

Cyprus
Panayiotis Zaphiris
Cyprus Computer Society

Czech Republic
Zdeněk Míkovec
Czech Society for Cybernetics & Informatics

Denmark
Torkil Clemmensen
Danish Federation for Information Processing

Finland
Kari-Jouko Räihä
Finnish Information Processing Association

France
Philippe Palanque
Société des Electriciens et des Electroniciens (SEE)

Germany
Tom Gross
Gesellschaft für Informatik

Hungary
Cecilia Sik Lanyi
John V. Neumann Computer Society

Iceland
Marta Kristin Larusdottir
The Icelandic Society for Information Processing

India
Anirudha Joshi
Computer Society of India

Ireland
Liam J. Bannon
Irish Computer Society

Italy
Fabio Paternò
Italian Computer Society

Japan
Yoshifumi Kitamura
Information Processing Society of Japan

Korea
Gerry Kim
KIISE

Malaysia
Chui Yin Wong
Malaysian National Computer Confederation

Netherlands
Vanessa Evers
Nederlands Genootschap voor Informatica

New Zealand
Mark Apperley
New Zealand Computer Society

Nigeria
Chris C. Nwannenna
Nigeria Computer Society

Norway
Dag Svanes
Norwegian Computer Society

Poland
Marcin Sikorski
Poland Academy of Sciences

Portugal
Pedro Campos
Associacão Portuguesa para o Desenvolvimento da Sociedade da Informação

Slovakia
Vanda Benešová
The Slovak Society for Computer Science

South Africa
Janet L. Wesson
The Computer Society of South Africa

Spain
Julio Abascal
Asociación de Técnicos de Informática

Sweden
Jan Gulliksen
Swedish Computer Society

Switzerland
Solange Ghernaouti
Swiss Federation for Information Processing

Tunisia
Mona Laroussi
Ecole Supérieure des Communications De Tunis

United Kingdom
Andy Dearden
British Computer Society

United States of America
Gerrit van der Veer
Association for Computing Machinery

Expert Members

Nikos Avouris (Greece)
Simone D. J. Barbosa (Brazil)
Peter Forbrig (Germany)
Joaquim Jorge (Portugal)
Paula Kotzé (South Africa)
Masaaki Kurosu (Japan)
Gitte Lindgaard (Australia)
Zhengjie Liu (China)
Fernando Loizides (Cyprus)
Dan Orwa (Kenia)
Frank Vetere (Australia

Working Group Chairpersons

WG13.1 (Education in HCI and HCI Curricula)
Konrad Baumann, Austria

WG13.2 (Methodologies for User-Centered System Design)
Marco Winckler, France

WG13.3 (HCI and Disability)
Helen Petrie, UK

WG13.4 (also 2.7) (User Interface Engineering)
Jürgen Ziegler, Germany

WG13.5 (Resilience, Reliability, Safety and Human Error in System Development)
Chris Johnson, UK

WG13.6 (Human-Work Interaction Design)
Pedro Campos, Portugal

WG13.7 (HCI and Visualization)
Achim Ebert, Germany

WG 13.8 (Interaction Design and International Development)
José Adbelnour Nocera, UK

WG 13.9 (Interaction Design and Children)
Janet Read, UK

Conference Committee

General Conference Co-Chairs
Tom Gross, Germany
Julio Abascal, Spain

Full Papers Co-Chairs
Simone D. J. Barbosa, Brazil
Philippe Palanque, France

Short Papers Co-Chairs
Fabio Paterno, Italy
Kari-Jouko Räihä, Finland

Posters and Demos Co-Chairs
Stephen Brewster, UK
David McGookin, UK

Organisation Overviews Co-Chairs
Melanie Fitzgerald, United States of America
Kori Inkpen, United States of America

Panels Co-Chairs
Anirudha N Joshi, India
Gitte Lindgaard, Australia

Tutorials Co-Chairs
Christoph Beckmann, Germany
Regina Bernhaupt, France

Workshops Co-Chairs
Christoph Beckmann, Germany
Víctor López-Jaquero, Spain

Doctoral Consortium Co-Chairs
Geraldine Fitzpatrick, Austria
Panayiotis Zaphiris, Cyprus

Proceedings Chair
Marco Winckler, France

Madness Co-Chairs
Artur Lugmayr, Finland
Björn Stockleben, Germany
Tim Merritt, Denmark

Local Organisation Co-Chairs
Mirko Fetter, Germany
Claudia Tischler, Germany

Student Volunteers Co-Chairs
Robert Beaton, United States of America
Sascha Herr, Germany

Contents

Doctoral Consortium 1

Face-to-Face Applications: the Usage of Large Screens to Increase Co-Present Social Engagement 3
Alessio Bellino

An Investigation into Freeform, Dynamic, Digital Annotation for Understanding Program Code 9
Craig Sutherland, Andrew Luxton-Reilly, Beryl Plimmer

Tools and Methods for Supporting the Development of Ubiquitous End-User Simulations 15
Miriam Greis

Temporal Transitions of User Experience 21
Daniela Wurhofer

Designing for Collaborative Sensemaking: Leveraging Human Cognition For Complex Tasks 27
Participant: Nitesh Goyal, Supervisor: Susan R. Fussell

Understanding Collaborative Activities: A Distributed Cognition Perspective 33
Christina Vasiliou

Accessibility for Cognitive and Mentally Disabled Persons 39
Stefan Johansson

Multimodal Approaches for Text Entry in Indian Language on Mobile Devices 43
Sanjay Ghosh

A Multi-Modal System for Public Speaking 49
[1]Fiona Dermody [2]Dr. Alistair Sutherland [3]Dr. Margaret Farren [2,3](Supervisors)

Improving Accessibility Support for Web Developers 55
David Swallow

Towards a Methodology to Evaluate Multimodal Games for Cognition in People who are Blind 61
Ticianne Darin

Bridging the Digital Gap in South Africa: Fitts's Model Impact on Usability 67
Guy Toko, Ernest Mnkandla

A Model for Parental Control Systems on Mobile Devices 73
Ibrahim R. Mbaya

Workshop: IFIP WG 13.5 Workshop on Resilience, Reliability, Safety and Human Error in System Development 79

Alignment of Technology to Work: Design & Evaluation Representation 81
Dorrit Billman[1], John Archdeacon[2], Rohit Deshmukh[1], Michael Feary[3], Jon Holbrook[1], Michael Stewart[1]

Reflecting on Users' Strategies for Resilient Interactions 95
Jonathan Day, George Buchanan, Stephann Makri

Accounting for Organisational faults in Task Model Based Systematic Analysis of System Failures and Human Errors 101
Camille Fayollas, Célia Martinie, Philippe Palanque, Racim Fahssi

Modeling Monitoring Behavior for HMI Designs is Easy with the Right Tool 119
Bertram Wortelen[1], Sebastian Feuerstack[2], Marcus Behrendt[2]

Workshop: Human Work Interaction Design (HWID): Design for Challenging Work Environments 127

Interaction Design for Stratigraphic Analysis in Archaeology 129
Barbara Rita Barricelli, Stefano Valtolina

Into The Woods 137
Judy Bowen, Annika Hinze, Sally Jo Cunningham

Designing for the Factory: UX Prototyping for the Cleanroom 145
Roland Buchner, Verena Fuchsberger, Astrid Weiss, Manfred Tscheligi

Experience Design and Positive Design as an alternative to classical human factors approaches 153
Michael Burmester[1], Katharina Maria Zeiner[1], Magdalena Laib[1], Cristina Hermosa Perrino[1], Marie-Luise Queßeleit[2]

Challenges for Action Research on HWID in Activity Based Workplaces 161
Åsa Cajander[1], Gerolf Nauwerck[1], Thomas Lind[1], Marta Larusdottir[2]

The Form of HWID Theory 171
Torkil Clemmensen

Understanding UI Design for Creative Writing: A Pilot Evaluation 179
Frederica Gonçalves[1], Pedro Campos[1], Anant Garg[2]

Pen and Display: A Multimodal Interaction Approach for Older Office Employees 187
Georg Regal[1], Ulrich Lehner[1], Valentin Gattol[1], Jan Bobeth[1], Manfred Tscheligi[1,2]

Out in the Cold, the Loneliness of Working with Doctors and Patients 195
Bert Vandenberghe, David Geerts

Workshop: IFIP WG 13.2 Workshop on User Experience and User-Centered Development Processes 201

User Experience Centered Engineering: A Process Model Inspired by Games Development 203
Regina Bernhaupt, François Manciet* and Michael Pirker***

Mining Logs to Support HCI (Re)Design 211
Felipe Cordeiro de Paula, Simone D.J. Barbosa

Integrating Human-Centered and Model-Driven Methods in Agile UI Development 215
Holger Fischer, Enes Yigitbas, Stefan Sauer

Managing the Agile Process of Human-Centred Design and Software Development 223
Peter Forbrig[1] and Michael Herczeg[2]

Continuous User Experience Development 233
Kati Kuusinen

Addressing Usability and UX in Call for Tender for IT Products 239
Rosa Lanzilotti, Maria Francesca Costabile, Carmelo Ardito

User and Client Satisfaction in Agile Development 249
Marta Larusdottir[1], Effie Law[2], Åsa Cajander[3]

Engineering for User Experience: An Interactive TV Case Study 259
Michael M. Pirker [1], Regina Bernhaupt [1,2], François Manciet [2]

A Review of Milestones in the History of GUI Prototyping Tools 267
Thiago R. Silva, Jean-Luc Hak, Marco Winckler

Workshop: IFIP WG 13.7 Workshop on Designing Interaction and Visualization for Mobile Applications (DIViM 2015) 281

The Future of Visual Perception 283
Gerrit C. van der Veer

3DIM: An Interactive 3D Map to Visualize Geo-Spatial Data in Mobile Devices 293
Ragaad AlTarawneh, Carl S. Marshall, Selvakumar Panneer, Cindy K. Chung

Simplifying the Input of Perceived Exertion in the Mobile Context using Prediction 301
Janko Timmermann[1], Alexander Schiotka[2], Wilko Heuten[1], Susanne Boll[2]

Smart Ecosystems and the Impact on Mobile Interaction Design Methods 313
Claudia Nass, Konstantin Holl, and Christian Müller

Towards Optimizing the Sunburst Visualization for Smart Mobile Devices 323
Ragaad AlTarawneh[1], Shah Rukh Humayoun[1], Abdel-karim Al-Jaafreh[2]

Meanings of a Blurred Mobile-Home Context for People aged 50plus....335
Thomas Meneweger[1], Marianna Obrist[2], Manfred Tscheligi[1]

Workshop: First International Workshop on Interactivity in Healthcare Systems (IHS 2015) 343

Context Awareness in Communication around Fall Handling with PERS....345
Jan Van den Bergh[1], Shirley Elprama[2], Jasmien Decancq[2], An Jacobs[2], Karin Coninx[1]

Worlds Apart - Doctors' Technological Frames and Online Medical Records ...357
Åsa Cajander[1], Christiane Grünloh[2,3], Hanife Rexhepi[4]

Self-Help Obesity Prevention Program in Stokvels: A Social Media Intervention 369
W. Douglas Evans[1], Nelia P. Steyn[2], Marjanne Senekal[2]

MyData Collection for Personal Health: Concept Design of a Lifestyle App for Junior Athletes 381
Jonna Häkkilä[1], Mira Alhonsuo[1], Juho Rantakari[2], Ashley Colley[2], Lasse Virtanen[2]

The Future of Digitally Enabled Health Coaching – A Proposed Model 387
Charalampos Kyfonidis, Marilyn McGee-Lennon

Rome Wasn't Reached in a Day: How to Motivate Patients to Keep Walking? 399
Bert Vandenberghe[1], Jasper Vanhoof[2], Fabienne Dobbels[2], David Geerts[1]

Workshop: The Landscape of UX Requirements Practices 405

UX Requirements to Public Systems for All: Formalisation or Innovation 407
Jane Billestrup, Anders Bruun and Jan Stage

User Experience in Technically Complex and Safety Critical Work Domains: Two Case Studies in Aerospace and Aviation 429
[1]Dorrit Billman, [1]Jessica Lee, [2]Emilie Roth

A Multi-Method Approach to UX Requirements: Adapting to Agile & Lean Development 449
Gregorio Convertino

Requirements Elicitation for New Video Game Development Tools: A Case Study 465
Christos Fidas[1,2], Nikolaos Avouris[1], Ivan Orvieto[3]

Solution-based Requirements Capture with PDot in an E-Learning Context 485
Matthias Heintz, Effie Law

Guidelines to Specify HCD Activities in the Call for Tender for Public Administration Websites 497
Simon Mastrangelo[1], Rosa Lanzilotti[2], Maurizio Boscarol[3], Carmelo Ardito[2]

Design Requirements for Web Applications to Affect the End User Emotional State 507
Giulio Mori, Fabio Paternò, Ferdinando Furci

From Usability Workarounds to Usability Around Work 513
Kimmo Tarkkanen, Ville Harkke

User Experience Goals as a Guiding Light in Design and Development - Early Findings 521
Heli Väätäjä[1], Paula Savioja[2], Virpi Roto[3], Thomas Olsson[1] and Jari Varsaluoma[1]

Improving User Experience through Task Design and Evaluation Metrics in Research Projects 529
Xiaojun Yuan

Workshop: Fostering Smart Energy Applications (FSEA 2015) 537

IdleWars: a Pervasive Game to Promote Sustainable Behaviour in the Workplace 539

Evangelos Tolias[1], Enrico Costanza[1], Alex Rogers[1], Benjamin Bedwell[2] Nick Banks[3]

Value Sensitive Design Approach to Influence Energy-use Behaviour 547
Rachel Burrows, Peter Johnson, Hilary Johnson

Visualizing and Gamifying Water and Energy Consumption for Behavior Change.....555
Isabel Micheel[1], Jasminko Novak[12], Piero Fraternali[3], Giorgia Baroffio[3], Andrea Castelletti[3], Andrea-Emilio Rizzoli[4]

Promoting Energy-Efficient Behavior by Recommendations based on Energy Cultures.....565
Stephan Hammer, Fabian Segmüller*, Birgit Lugrin[+], Elisabeth André**

Towards using Low-Cost Opportunistic Energy Sensing for Promoting Energy Conservation 575
Nuno J. Nunes, Lucas Pereira, Valentina Nisi

Private Focus Portals to Shared Energy Visualizations 585
Chi Tai Dang[1], Masood Masoodian[2], Elisabeth André[1]

Presentation Methods to Inform Decisions about Energy Usage.....595
Chris Killeen

Watt-I-See: Probing Future Distributed Energy Scenarios 605
Clinton Jorge, Filipe Quintal, Nuno J. Nunes, Valentina Nisi

Illustrating Energy Related Properties of Buildings Using a 3D-Game-Engine 615
Thomas Rist, Jens Müller

Simulating the Impact of Household Energy Consumption on the Electricity Grid 623
Patrick Ozoh, Mark Apperley

Interactive Solar Panel Simulation Tool - From GHI to PV Output.....633
Joris Suppers, Mark Apperley

Doctoral Consortium

Face-to-Face Applications: the Usage of Large Screens to Increase Co-Present Social Engagement

Alessio Bellino
DISCo, University of Milano-Bicocca
Viale Sarca 336/14, Milan, Italy.
bellino@disco.unimib.it

Abstract. Together with Giorgio De Michelis and Flavio De Paoli (my PhD Supervisors, [gdemich, depaoli]@disco.unimib.it), we are investigating how to stimulate social engagement of small groups (e.g. best friends, couples) through shared applications displayed on large screens controlled by smartphones. The research area is IxD.

1 Introduction

In the last years, there has been the rapid worldwide diffusion of smartphones and tablets. Together with personal computers, they offer wide possibilities for interacting with digital content in any situation: (i) sitting at a desktop and (ii) standing or sitting anywhere. Moreover, the growth of large high-resolution displays allows the design of new systems offering new interaction possibilities. For example, [5] and [7] offer technical solutions to support simultaneous interaction of co-located people by sharing digital information on multi-touch large screens. Their solution is effective for small groups of people standing at arm's distance from the screen. At any rate, there could be ergonomic problems: the so-called 'Gorilla Arm syndrome' and the need of incessant moving when people interact on very large touch displays [3]. Moreover, in situations in which users involved prefer to interact in a more comfortable position, such a physical setting is not adequate for many reasons. For example, standing in front of the screen reduces the screen view to others participants (e.g., to students in a classroom sitting at their desks); people may prefer not to move to get close to the screen to touch it (e.g. family members would like to control the smart TV while sitting on the sofa in the living room). In similar cases, there is the need for a solution that lets users interact with the screen from a distance. Systems composed of a large screen and hand-held devices (e.g., smartphones) can overcome these limitations allowing users to interact with the digital content displayed by distant large screens. In these cases, there are interesting challenges for interaction designers, who

have to define (i) the correspondence between interfaces on the smartphone and the ones on the shared screen facilitating the command and control by users, and (ii) how simultaneous interaction of multiple users can be managed to avoid inconsistencies and conflicts.

The use of distant large screens to share experiences gets the most positive feedback in terms of potentiality and possibility of practical use within groups of co-located people [2]. Accordingly, our aim is to stimulate the social engagement of small groups through co-present interaction with social applications displayed on shared large screens.

2 Motivation: Base Scenario

Let us consider two friends (Alice and Bob) sharing a traditional application with standard control devices (e.g., keyboard and mouse). Only one user at a time can operate the application (Figure 1a): while Alice is free to act, Bob can only see Alice's behavior and what is displayed on the shared screen. Bob can be part of Alice's interaction by pointing at the screen to show the content of his interest, or asking Alice to perform a certain action.

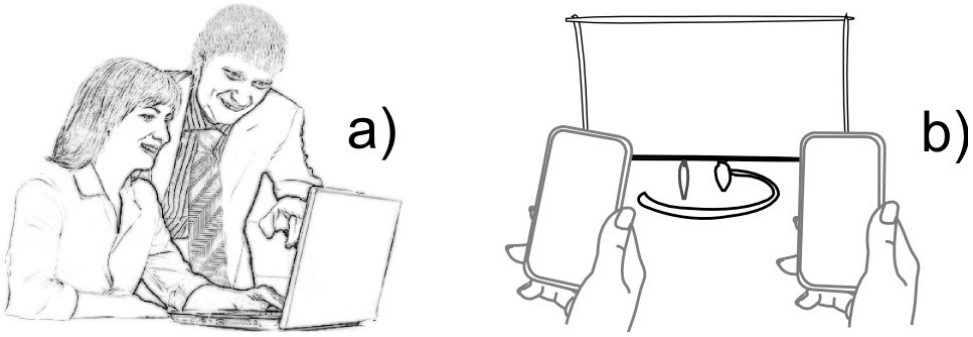


Figure 1: Alice and Bob using (a) traditional interaction and (b) social interaction.

To overcome these limitations, we let Alice and Bob use their smartphones as control devices removing the obstacle of one user at a time (Figure 1b): Alice and Bob can sit on a sofa comfortably, and interact with the shared screen, simultaneously, with the same possibilities (only when their actions interfere, the system may enforce lock-up mechanisms).

3 Overcome Challenges

3.1 Interacting with Large Screens: Smartphone Interaction

Different techniques to control distant screen have been investigated: pointer-based [6], direct (e.g. touch), bodily (e.g. Kinect) and mobile interaction (e.g. smartphones) [4]. According to the scenarios described in section 1, the most suitable technique seems to be the mobile interaction with some improvements. The existent techniques do not satisfy some of the principles we consider crucial: (i) users should be able to operate the smartphone with one hand, using the most common touch gestures with the thumb (drag/swipe and tap), without looking at it too often (the gaze should always be at what is displayed on the shared screen [4]); and (ii) the effects of control operations have to be reported on the visualization interface to give feedback making users aware of what is going on.

According to the first principle, the *action* is made on the smartphone and the corresponding effects are *visualized* on the large screen: *action* and *visualization* are decoupled. According to the second principle, different colors – one for each user – give feedback about what happens on the large screen.

[1] displays the presented interaction style. Moreover, a single-user prototype was designed to evaluate the effectiveness and naturalness of the interaction. It allows basic interactions such as the control of a set of pictures, audio and video streams, navigation and selection of elements, map control and so on. Users took a questionnaire using a 6-point Likert scale. Answers were given considering several factors including usability (Mean: 5.1, Mode: 5) and naturalness (Mean: 4.9, Mode: 5). Users generally appreciated the proposed interaction style.

3.2 Multi-User Simultaneous Interaction

Two factors could increase social engagement: (i) freedom of acting so that every user can take the initiative and (ii) social awareness so that every user can understand what the others are doing exploiting feedback on the shared screen. Therefore, how to manage concurrency?

When users interact with different areas of the large screen, there is no problem. Problems may arise when multiple users are acting on the same section since each of them may focus on a specific portion of the displayed content. For example, if two users are browsing the same list, a user may be focusing on the head of the list and the other on the tail. In such cases, it is

necessary to adopt presentation artifacts letting users co-navigate the list. A solution is provided by a multi-focus fisheye effect [1] to magnify the parts of the list associated with each user. Moreover, the border of each fisheye lens is colored (feedback aforementioned) to let watchers understand who is associated with that lens. In any other case in which there could be conflicts, two approaches can be adopted: (i) instruct the application to enforce control policies, or (ii) let users discuss and negotiate control policies that are then enforced by themselves (hence without inclusion of control mechanisms in the application). The former is preferred to prevent from unwanted situations (e.g., displaying sensible information without explicit consensus of the participants); the latter is advisable in the most common situations when users face many possible alternatives (e.g., select the next video to play among a shared set). Social awareness gives users the tools to make informed decisions engaging them.

4 Why a Doctoral Consortium?

Several questions are unanswered:

1) What kinds of face-to-face applications could be designed? YouTube4Two [1] is an example, but interactions are infrequent (users prefer not to interact while the video is playing).

We thought about a sort of “Facebook4Two” or an application to book hotels by two co-present persons. Gathering ideas at the consortium could be interesting.

2) How to measure social engagement given by applications designed for co-present use? How to measure “engagement” qualities of the proposed interaction?

3) How to go beyond smartphone interaction? Kinect or Nintendo Wii offer different interaction settings, but can they be used comfortably for controlling large screens? For example, users might suffer the ‘Gorilla Arm syndrome’ or get annoyed due to the ambiguity between wanted and unwanted commands.

References

[1] Bellino A. et al. YouTube4Two: socializing video platform for two co-present people. In Proc. of the companion publication of the 23rd int. conf. on World Wide Web companion. IW3C2.

[2] Jarusriboonchai, P., &Väänänen-Vainio-Mattila, K. How can technology bring families together: Exploring user needs and design qualities. In Proc. of the 13th Int. Conf. on CHI with Mobile Devices and Services, ACM (2011).
[3] Jurmu M. et al. Waving to a touch interface: descriptive field study of a multipurpose multimodal public display. In Proc. of the 2nd ACM International Symposium on Pervasive Displays, pages 7-12. ACM, 2013.
[4] Kurdyukova E. et al. Direct, bodily or mobile interaction?: comparing interaction techniques for personalized public displays. In Proc. Of the 11th Int. Conf. on Mobile and Ubiquitous Multimedia, page 44. ACM, 2012.
[5] Multitaction by multitouch (multitaction.com), March 2015.
[6] Nancel M. et al. High-precision pointing on large wall displays using small handheld devices. In Proc. of the SIGCHI Conference on Human Factors in Computing Systems, pages 831-840. ACM, 2013.
[7] Smart technologies (smarttech.com), March 2015.

An Investigation into Freeform, Dynamic, Digital Annotation for Understanding Program Code

Craig Sutherland, Andrew Luxton-Reilly, Beryl Plimmer
Department of Computer Science
University of Auckland
New Zealand
{cj.sutherland|a.luxton-reilly|b.plimmer}@auckland.ac.nz

Abstract. Understanding program code is a time-consuming task. Previous research has found freeform annotations to be an invaluable tool in assisting comprehension for other forms of prose. Therefore freeform annotations may also increase code comprehension. However code has some unique characteristics compared to other forms of prose. My research is investigating how we can implement freeform annotations inside a code editor in a way that will reduce the time programmers spend on understanding code.

1 Research Area and Topic

My research area is stylus input computing and digital ink annotations.

I am investigating how freeform, digital ink can assist programmers comprehend computer program code and how freeform annotations should automatically adapt when the context changes.

2 Research Problem

Programmers spend a lot of time understanding computer program code. Program code differs from other forms of prose: it is non-linear, highly structured and dynamic. These differences increase the cognitive demands [1].

Programmers need to remember complex paths through the program code. This requires remembering what they have previously read and being able to quickly navigate through the code. Indeed, integrated development environments provide many tools that implement these two tasks. But these tools are infrequently utilised; partly due to the learning curve required [2].

Annotations assist comprehension by reducing the cognitive workload. They allow the reader to offload information from their short term memory and streamline navigation through a document [3]. This suggests annotations

could help reduce the effort required by programmers when reading program code.

The problem with current program annotation tools is they are text-based: forcing a reader to type text increases their cognitive workload [4]. In contrast, hand writing annotations does not appear to increase cognitive workload [4]. Therefore, if a programmer could use hand-written annotations on their code they would realise the benefits of annotations. However very few programmers print computer code because its size and dynamic nature. Instead programmers read code on-screen which limits their ability to hand-write annotations.

There are now tools that allow programmers to annotate code on-screen within their development environment (e.g. [1]). These tools proved that it is possible to annotate code on-screen but they also identified several other challenges.

One challenge is how to maintain the meaning of annotations. Programmers often want to understand code they are changing. If they annotate some code and then change the code then the annotations lose their meaning. Ideally annotations should adapt to changes in the content but people are unused to seeing their annotations change. Research is needed to identify when and how annotations should adapt in response to changes in the underlying content.

A second challenge is how to integrate annotations in a way that enhances comprehension. Previous research identified benefits of freeform annotations in digital readers for comprehension [5]. Further investigation is needed to see if these benefits also apply to program code.

3 Research Claim

Freeform, dynamic, digital ink annotations will reduce the time programmers need when trying to understand computer code.

4 Methods

The first phase of my research was a user study investigating how programmers annotate program code on paper. I observed experienced programmers reading and annotating code and then interviewed them asking why they annotated and the significance of their annotations. These findings have helped formulate the requirements of an annotation system for programmers. This study is to be reported at INTERACT 2015 [6].

The second phase is testing different ways of automatically changing annotations when the underlying code changes. The literature review identified some ways that annotations could be modified but these are very limited. In this study the participants read and annotate some code in an extension to Visual Studio. When they finish I modify the code underneath their annotations so the annotations are adapted. The participants rank the different automatic adaption routines based on preserving the original meaning. This is followed by an interview on what they think of the different adaptions.

The final phase will use the results from the previous phases to develop and evaluate an annotation system for programmers. The first phase provided the background on programmers' annotation practices. These will be used for building the workflow to support annotation. The second phase will provide an implementation of dynamic annotations. The aim of the final phase is a system that provides complete end-to-end support that assists program code comprehension. This solution will be validated in a field trial with software engineering students.

5 Proposed Solution

The final proposed solution will contain two parts: algorithms for automatically adapting annotations in response to code changes; and tools to assist programmers in navigating the code.

The algorithms use the type of annotation and the underlying code so the original meaning of the annotations is preserved.

The navigation tools will assist the programmer in both re-finding previous code and serendipitously finding new code. These will work without increasing the cognitive workload of the programmer.

6 Expected Contributions

A systematic literature review on how digital ink annotations have been implemented previously.

- Some guidelines on how to automatically adapt annotations in response to context changes.
- Formal evaluations of:
  - User expectations around how users react to automatic changes in their annotations and;
  - How annotations can support code comprehension.

7 Statement of Work

To date I have completed the systematic literature review. The review has been submitted to a journal for publication and we are waiting for the feedback. This study identified gaps around user expectations of automatic adaptations and how freeform annotations can be used to assist with software comprehension.

The first user study has been completed and accepted for publication at INTERACT 2015 [6].

The second user study has ethics approval and I have started collecting data. The prototype has been implemented allowing programmers to annotate code in Visual Studio. Annotations are automatically adapted using six different algorithms based on the type of annotation.

I have started implementing the functionality for the final iteration. This is based on requirements gathered from the first user study. However I have still to decide exactly which functionality to implement and how it should be tested.

My current issues are exactly what functionality would be beneficial for reading comprehension and how to assess the effectiveness of the functionality. I am thinking of doing two to three rounds of testing. After each round I will tweak the functionality to try and improve comprehension. These are the issues I would like to discuss with the consortium to gain some feedback on the final implementation.

References

[1] Priest, R. and Plimmer, B., "RCA: experiences with an IDE annotation tool," in Proceedings of the 7th ACM SIGCHI New Zealand chapter's international conference on Computer-human interaction: design centered HCI, Christchurch, New Zealand, 2006, pp. 53-60.

[2] Maalej, W., Tiarks, R., Roehm, T., and Koschke, R., "On the Comprehension of Program Comprehension," ACM Trans. Softw. Eng. Methodol., vol. 23, pp. 1-37, 2014.

[3] Crisp, V. and Johnson, M., "The use of annotations in examination marking: opening a window into markers' minds," British Educational Research Journal, vol. 33, pp. 943-961, 2007.

[4] O'Hara, K. and Sellen, A., "A comparison of reading paper and on-line documents," in Proceedings of the ACM SIGCHI Conference on Human factors in computing systems, Atlanta, Georgia, USA, 1997, pp. 335-342.

[5] Schilit, B. N., Golovchinsky, G., and Price, M. N., "Beyond paper: supporting active reading with free form digital ink annotations," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Los Angeles, California, USA, 1998, pp. 249-256.
[6] Sutherland, C. J., Luxton-Reilly, A., and Plimmer, B., "An Observational Study of How Experienced Programmers Annotate Program Code," in INTERACT 2015, Bambery, Germany, 2015.

Tools and Methods for Supporting the Development of Ubiquitous End-User Simulations

Miriam Greis
University of Stuttgart
Institute for Visualization and Interactive Systems
Pfaffenwaldring 5a, 70569 Stuttgart
miriam.greis@vis.uni-stuttgart.de

Abstract. Predictive simulation is a powerful technique used by experts to support them in decision-making. We envision that non-experts will also use this technique in the future. Interfaces showing predictions, as, e.g., navigation systems that predict the arrival time, are already on the market. An important aspect of these interfaces is the uncertainty of the input parameters and the output. In our research, we aim at developing interactive input methods and visualizations for communicating uncertain data. Therefore we develop a modelling and simulation tool for non-experts which allows to evaluate interactive input methods for uncertain parameters and different types of visualization for uncertain data. The results of our research will support the future development of predictive simulations for non-experts and help to avoid unintentional manipulation by choosing the best input methods and visualizations.

1 Introduction

Simulation is a powerful technique to build and examine models of real-world systems. It is one of the most used techniques in research and management science and can be applied in many disciplines [4]. Simulations help to better understand or develop strategies for the management of complex systems in cases where adapting the real system is impossible or too expensive [8].

At the moment, simulations are only used by experts that have specialized knowledge in model building or simulation execution and both programming and mathematical skills. So far, non-experts, which we consider to be people without programming skills and knowledge about simulations are not able to use simulations, although simulations could also support decision-making in everyday life. Nevertheless already today, people rely on predictions about, e.g., the weather or the arrival time when traveling. They additionally use predictions implicitly to foresee future events and plan activities. Predictive simulation would be a powerful technique to support non-experts in making predictions explicit and more reliable.

One key aspect of simulation usage is the input and output of uncertainty. During the modeling process, uncertainty in the assumptions of the model and uncertainty in the parameter choice have to be taken in account. Thus, also

simulation results are uncertain. Allowing the input of parameter uncertainty and communicating the uncertainty in the output is crucial to avoid unintentional manipulation of non-expert users.

The research supervised by Prof. Dr. Albrecht Schmidt therefore aims at developing and evaluating input and output methods for uncertain data in the context of predictive simulations.

2 Related Work

Simulation tools for non-experts with an educational purpose were already developed very early. Two of the first tools were the Alternate Reality Kit [10] and Playground [2]. Smith et al. [9] and Cypher et al. [1] developed the simulation tool KidSim especially for children. It uses the principle of programming by demonstration/example. StarLogo [6] allows the user to attach puzzle pieces to each other instead of programming to construct a model. Another simulation tool for non-experts is NetLogo [11], which requires users to have a basic programming knowledge for creating own models. All these tools use agent-based modeling to impart knowledge about simulations, but could not be used to support decision-making for non-experts in everyday life.

No work so far was done about the input of parameter uncertainty for non-experts, but a lot of previous work exists on the output of uncertain information. One strand of work for example investigated uncertainty information in weather forecasts. Morss et al. [5] found that most people are aware of the uncertainty in deterministic weather forecasts, although the range of this uncertainty was perceived very differently. Additionally, 70 % of the people preferred forecasts that contained information about the uncertainty of the forecast. Studies by Roulston et al. [7] and Joslyn et al. [3] showed that people make better decisions when having information about the uncertainty of a forecast and that information about the uncertainty also increases the trust in a forecast.

Overall, previous work about uncertainty communication focused on very specific scenarios from which results cannot be generalized. Additionally, there was no research-in-the-large approach used so far.

3 Research Approach

Our research approach includes three main aspects. First of all, we investigated strategies and tools used by simulation experts to get an overview about simulation usage and challenges. Second, we started to analyze the usage behavior of non-experts to understand how they currently use predictions and simulation results. Additionally, we identified challenges, benefits, and wishes to develop use cases and requirements for the future

usage of simulation services. The third part of our research approach includes the design and implementation of a simulation tool that allows domain experts to create models that can then be executed by non-experts.

3.1 Research Claim

We assume that non-experts will be able to understand basic concepts of modeling and simulations and can use easy tools to interact with them. This includes the input and understandable output of uncertain data. Input and output modalities for uncertain data can be standardized and categorized to make the input and output of uncertainty easier for users and developers.

3.2 Methods

At the beginning of our work, we used multiple different methods to get insights into current simulation usage of experts and non-experts. We conducted an online survey and a paper questionnaire with researchers working with simulations. We also run a diary study with non-experts to understand how they already use predictive simulations in everyday life and conducted focus groups with both experts from research and non-experts to get insights in their understanding of the term simulation and potential future usage scenarios. We additionally conducted an online survey to compare existing representations that involve uncertainty.

Currently, we plan to do research-in-the-large with an online game and a web-based online modeling tool. With the help of the modeling tool, we plan to investigate a specific usage scenario and evaluate whether non-experts are able to construct valid models with the help of an interactive platform. We additionally will use the web-based tool to run user studies on the usage behavior and user experience of input and output methods for uncertainty for methods that already exist and for methods that we specifically designed. We plan to further evaluate our input and output methods with the help of eye tracking.

3.3 Solution and Contributions

The research work contributes a web-based framework that allows the evaluation of input and output methods for modeling and simulation tasks and will as well support the collection of usage data. The included modeling tool allows non-experts to model relationships for at least one specific use case. The research further contributes ideas, use cases, and prototypes for new simulation services that can support non-experts in everyday life.

With the help of the web-based simulation framework, we will also provide a set of evaluated input methods and a set of evaluated output methods for uncertain data that will be made available for web developers.

4 Open Questions and Issues

Open questions include sharpening the research topic and finding a focus. Additionally, finding connections to current topics in Human Computer Interaction, other researchers, and the community in general is an open issue.

Acknowledgements

The author would like to thank the German Research Foundation (DFG) for financial support of the project within the Cluster of Excellence in Simulation Technology (EXC 310/2) at the University of Stuttgart.

References

[1] A. Cypher and D. C. Smith, “KidSim: end user programming of simulations,” Proc. SIGCHI Conf. Hum. factors Comput. Syst., pp. 27–34, 1995.
[2] J. Fenton and K. Beck, “Playground: an object-oriented simulation system with agent rules for children of all ages,” ACM SIGPLAN Notices, vol. 24. pp. 123–137, 1989.
[3] S. L. Joslyn and J. E. LeClerc, “Uncertainty forecasts improve weather-related decisions and attenuate the effects of forecast error.,” J. Exp. Psychol. Appl., vol. 18, no. 1, p. 126, 2012.
[4] A. Law and W. Kelton, “Simulation modeling and analysis,” Journal of Manufacturing Systems, vol. 18. McGraw-Hill, p. 382, 1999.
[5] R. E. Morss, J. L. Demuth, and J. K. Lazo, “Communicating Uncertainty in Weather Forecasts: A Survey of the U.S. Public,” Weather Forecast., vol. 23, no. 5, pp. 974–991, Oct. 2008.
[6] M. Resnick, “StarLogo: An Environment for Decentralized Modeling and Decentralized Thinking.,” in CHI 96 Conference companion on Human factors in computing systems, 1996, pp. 11–12.
[7] M. S. Roulston, G. E. Bolton, A. N. Kleit, and A. L. Sears-Collins, “A laboratory study of the benefits of including uncertainty information in weather forecasts,” Weather Forecast., vol. 21, no. 1, pp. 116–122, 2006.
[8] R. E. Shannon, “Introduction to the art and science of simulation,” 1998 Winter Simul. Conf. Proc., vol. 1, 1998.
[9] D. C. Smith, A. Cypher, and J. Spohrer, “KidSim: programming agents without a programming language,” Communications of the ACM, vol. 37. pp. 54–67, 1994.

[10] R. B. Smith, “The Alternate Reality Kit; An animated environment for creating interactive simulations.,” in IEEE Computer Society Workshop on Visual Languages, 1986, pp. 99–106.
[11] S. Tisue and U. Wilensky, “Netlogo: A simple environment for modeling complexity,” Int. Conf. Complex Syst., pp. 1–10, 2004.

Temporal Transitions of User Experience

Daniela Wurhofer

Center for Human-Computer Interaction, University of Salzburg
Sigmund-Haffner Gasse 18, 5020 Salzburg, Austria
daniela.wurhofer@sbg.ac.at
Advisor: Prof. Manfred Tscheligi
manfred.tscheligi@sbg.ac.at

Research area: User Experience, Temporality, Temporal Transitions.

Research topic. In my PhD thesis, I characterize temporal transitions of User Experience (UX) through theoretical and empirical exploration.

Research problem. Research on the temporality of UX has received increased attention within the HCI community over the recent years (e.g., [2], [5]), and different facets or notions of temporal UX have been studied. One such facet is the change of *specific UX factors* over time (e.g., perceived safety, stress, emotions), which has been investigated in different contexts (e.g., HRI [1], mobile phones [4]). Further, different *phases of UX* (i.e., anticipated, actual, or recalled experiences) have been investigated with regard to characteristics of specific phases as well as shifts or differences between the phases. For example, recalled experiences have been discussed by Kujala [5], or Karapanos [3], and the role of anticipated experiences has been investigated by e.g., Olsson [8] or Yogasara [10]. Shifts in UX over time have been investigated for the usage of mobile phones by Karapanos [2]. Besides, different *perspectives on UX* can be taken when investigating temporal UX (i.e., holistic vs. reductionistic perspective), shaping the research process and methods applied. Work on theoretical framings of temporal UX has been for example done by Pohlmeyer et al. [9] or Luojus [7].

Despite a range of empirical and theoretical work in the field of temporal UX, the gained insights are not interrelated to each other (i.e., standing alone), missing to frame temporal UX from a more systematic and broader perspective to guide future research in this area.

In my work, I define these different facets and notions of temporal UX as *temporal transitions of UX*, emphasizing temporal UX as a complex and multi-faceted phenomenon. To provide a comprehensive view on the

temporality of UX (including different facets and notions of temporal UX as well as their interrelation), it is necessary to theoretically frame and structure temporal transitions of UX and refine the theoretical framing based on empirical data. Thus, I create a more comprehensive picture of the temporality of UX, allowing to integrate previous insights and guide future research in this area.

Relevance of research. I argue that a more deep systematic theoretical and empirical analysis of temporal UX is needed. In particular, I want to address the following four dimensions: 1) theoretical framings of temporal UX require more empirical data and long-term studies as input [4], 2) to adequately reflect current research processes, there should be a combined holistic-reductionistic perspective on UX [6], 3) changes of specific UX factors in specific contexts need to be investigated systematically, 4) shifts and differences across different phases of UX (i.e., anticipated, actual, or recalled experiences) have to be examined in more detail. By addressing these issues in my work, knowledge on temporal UX as well as UX in general will be advanced.

The research hypothesis (claim). The overall objective of this work is to characterize temporal transitions of UX. I therefore propose the following research questions:

- RQ1: How to theoretically frame temporal transitions of UX?
- RQ2: How to empirically describe and characterize temporal transitions of UX?
  - RQ2a: How do specific UX factors change over time in specific contexts?
  - RQ2b: How can shifts and differences regarding UX phases (i.e., anticipated, actual and recalled UX) be characterized?

A sketch of the proposed solution. In order to answer the research questions, my approach can be divided into three main steps (see Figure 1).

In a first step, I aim at establishing an overall understanding of UX based on a meta analysis of existing literature and a perspective which combines both the holistic and the reductionistic view on UX. This serves as a structural basis for my further research (*RQ 1*).

In a second step, addressing *RQ2,* I focus on the empirical investigation of transitions of UX on two levels: On the one hand, I study changes of selected UX factors (e.g., perceived safety, stress) in two exemplary contexts (i.e., car and factory) (*RQ2a)*. On the other hand, I investigate shifts and differences across different phases of UX (i.e., anticipated, actual and recalled UX) as

well as differences between the three phases *(RQ2b)*. This will be also done in the two exemplary contexts. For the empirical studies, I apply a mix of qualitative (e.g., narrative interviews) and quantitative (e.g., questionnaires) methods.

In a third step, I will incorporate the empirical insights in the theoretical framework and iterate it based on the insights (*RQ 1*). Thus, I will create a framework which is based on both literature and empirical studies.

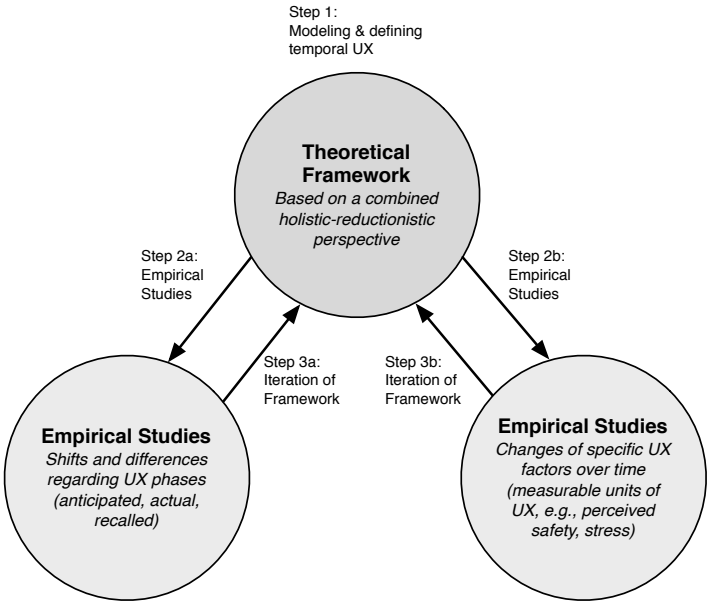


Figure 1. A sketch of the proposed solution.

The expected contributions of the PhD research. My PhD thesis can be split in a theoretical and an empirical contribution, which inform each other iteratively:

a) A theoretical framework describing temporal transitions of UX, including the different facets of temporal UX (e.g., shifts of UX, phases of UX).
b) Empirical data providing an improved understanding of temporal transitions of UX

This contributes to the overall objective of creating a more comprehensive picture of the temporality of UX including different facets (e.g., temporal phases of UX, changes of specific UX factors) and their interrelation.

A statement of work to date and open questions/issues for the discussion at the DC. So far, I set up a theoretical framework for UX (RQ1), providing a

structural basis for my research. To empirically investigate temporal transitions of UX (RQ2), initial studies in the factory and in the automotive context were performed. Based on the insights of these initial studies, further studies are set up and conducted. Regarding the participation in the doctoral consortium, I expect to get input of how the theoretical and the empirical part of my work can be interrelated optimally. I am curious to get feedback on my two application contexts and how they could be optimally used to provide insights regarding my research questions. Further, I would be happy to discuss how to theoretically model the multi-dimensionality of temporal UX.

Acknowledgements

The financial support by the Austrian Federal Ministry of Science, Research and Economy and the National Foundation for Research, Technology and Development is gratefully acknowledged (Christian Doppler Laboratory for “Contextual Interfaces”).

References

[1] Buchner, R., Wurhofer, D., Weiss, A. and Tscheligi, M. Robots in Time: How User Experience in Human-Robot Interaction Changes over Time. In ICSR International Conference on Social Robotics (2013).

[2] Karapanos, E., Zimmerman, J., Forlizzi, J. and Martens, J. (2009). User experience over time: an initial framework. In Proc. CHI '09. ACM, New York, NY, USA, 729-738.

[3] Karapanos, E. Martens, J. and Hassenzahl, M. (2010). On the retrospective assessment of users' experiences over time: memory or actuality? In CHI EA '10. ACM, New York, NY, USA, 729-738.

[4] Kujala, S. and Miron-Shatz, T. (2013). Emotions, experiences and usability in real-life mobile phone use. In Proc of CHI ’13, pages 1061–1070.

[5] Kujala, S., Roto, V., Väänänen-Vainio-Mattila, K., Karapanos, E. and Sinnelä, A. (2011). UX Curve: A method for evaluating long-term user experience. Interact. Comput. 23(5), 473-483.

[6] Law, E. (2011). The measurability and predictability of user experience. In Proc. EICS ’11. ACM, New York, NY, USA, 1–10.

[7] Luojus, S. (2010). From a momentary experience to a lasting one - The concept of and research on expanded user experience of mobile devices. PhD thesis, Acta Universitatis Ouluensis Scientiae Rerum Naturalium.

[8] Olsson, T. (2014). Layers of user expectations of future technologies: an early framework. In CHI EA '14. ACM, New York, NY, USA, 1957-1962.

[9] Pohlmeyer, A.E., Hecht, M., and Blessing, L. (2009). User experience lifecycle model continue [continuous user experience]. Der Mensch im Mittepunkt technischer Systeme. VDI Reihe, 22:314–317.

[10] Yogasara, T. (2014). Anticipated user experience in the early stages of product development. PhD thesis, Queensland University of Technology.

Designing for Collaborative Sensemaking: Leveraging Human Cognition For Complex Tasks

Participant: Nitesh Goyal, Supervisor: Susan R. Fussell
Cornell University
Dept. of Information Science
Ithaca, NY 14850, USA
ngoyal@cs.cornell.edu, sfussell@cornell.edu

Research Area. Collaborative Sensemaking, Crowdsourcing, User Interfaces for Web Applications

Research Topic. My research aims to design systems for complex sensemaking by remotely located non-expert collaborators (crowds), to solve computationally hard problems like crimes.

1 Research Problem and Related Work

While sensemaking has been studied in the past, designing interfaces for relatively complex sensemaking where experts and non-experts may collaborate remains a challenge. I am interested in how such collaborative sensemaking by crowds (non-experts) may help us better solve unstructured problems, where traditional computational techniques have failed. In particular, I am researching how to design for experts and non-experts (crowds) in the crime-solving domain, where solutions are often found through serendipity, instead of rules.

Crowdsourcing for somewhat complex tasks has been pursued in the past. Collaborative document editing in Soylent [2], creating taxonomy of colours in Cascade [3], suggesting a travel itinerary under constraints using Mobi [4], and mining sentiments by crowds for text analytics in OpinionBlocks [12] are some recent forays where crowdsourcing has shown to be performant and/or efficient. However, more open-ended domains like crime solving, requiring serendipitous discovery of clues and criminals, have yet to be crowd-sourced successfully.

As number of workers and associated workflows grow in complexity, crowdsourcing can be challenging [1]. Crowdsourcing for complex workflows has been pursued also. For example, CrowdForge by Kittur et al [5] explains how map-reduce framework popularized by Google may be used to partition

bigger complex tasks into smaller tasks dynamically by workers. Further, Malone et al 's [6] aggregation dimension suggests that crowd-workers can either work alone independently or depend upon each other to work together. Little et al [7]'s TurKit can further help decide what to present to each worker such that the flow of results of tasks between dependent workers can be controlled.

As such crowd-workflows become complex, researchers must identify the level of crowd-supervision needed for optimal output. Kulkarni et al [8] designed Turkomatic based on price-divide-loop such that real time visualization of the workflow-design is evident because unsupervised crowds failed to produce proper workflows resulting in a less than optimal output. On the other hand, supervised crowds in a conversational-agent, Chorus by Lasecki et al [9] made users believe that a single user exist behind Chorus. Instead, Chorus employs multiple crowd workers who collectively create response possibilities, such that Crowd workers can learn and remember collectively.

To summarize, while relatively complex tasks and workflows using crowds have been attempted, we have yet been unable to design a system that may structure non-experts (lesser trained crowds) and experts (trained workers) together in an interface to solve complex challenges like crime solving.

2 Research Hypothesis

The core aim of this research is to pursue a user centred design approach to designing a web-interface, and underlying system, that may enable collaborations between experts and non-experts, and within non-experts. I hypothesize that a user-centered-designed web-interface that supports collaboration between non-expert crowds for solving carefully broken down micro-tasks would help leverage distributed human crowd-cognition to solve complex tasks like crimes.

3 Proposed Method

I plan to integrate my findings based on a mixed-methods study. First, I will understand how trained-non-experts (trained students, through video and usage-log analysis) solve complex problems singularly and collaboratively. Consequently, I will extract important features that result in success and failure in problem solving. Based on these features, I propose to create a web-interface for collaborative problem solving. Finally, I will design a study to

validate whether the identified features lead to success or failure with non-expert crowds? So, based on the iterative nature of design process, my proposed solution would involve multiple iterations and steps before I design the final interface:

Step 1. Understand role of currently used features for solo sensemaking.

Step 2. Explore effects of information-sharing collaborative sensemaking.

Step 3. Extract features to identify micro-tasks, and workflows for success.

Step 4. Design web-interface for experts and non-experts to collaborate.

Step 5. Design a set of user-studies to measure user-experience, and performance achieved by non-experts with the web-interface at solving crimes.

4 Progress So-far

I have completed two iterations of system building of SAVANT tool to support solo [10] and collaborative sensemaking [11, 12] to better understand role of different design features:

In Iteration 1 (Step 1), I tested the utility of system-generated visualization of data links and a notepad for collecting annotations, and found system-generated visualizations to be significantly important in solving crimes [10].

In Iteration 2 (Step 2), I explored value of implicitly sharing insights by self-created visualizations of annotations, without explicitly pushed/requested information by collaborators. When implicit sharing of notes and self-created visualization of these notes was available, users identified more clues [11, 12].

5 Next Steps

I am presently working on Task Analysis for Step 3. I am conducting a qualitative video-analysis, and usage-log analysis of the actions performed by successful and unsuccessful pairs in Step 2. Based on video-analysis, 3 *design goals* seem promising for success: *externalizing insights; shoe-boxing visually*; *and iterating over previously collected information*. I am also identifying user-actions, based on interface-log analysis, when pursued multiple times by users would lead to successful resolution of the task.

Based on preliminary findings, I am designing the web-interface (SAVANT) for non-experts and experts with recommended steps associated with success. Based on these findings, I'd be better equipped with knowledge of micro-tasks, workflows, and atomicity of the dataset that would enable success in task-resolution. So, I'd propose using the lessons learnt to design

the next SAVANT version where users using the full SAVANT suite might be able to collaborate and auto-direct micro-tasks to crowds that would support/challenge their own insights and help resolve the crime-solving task.

Finally, I would like to use the INTERACT 2015 Doctoral Consortium to help design user-studies and determine measures to evaluate user-experience, and features of SAVANT like system-generated micro-tasks vs. manually generated micro-tasks; and presence of messaging with the crowds vs. no messaging.

6 Proposed Solution

My proposed solution would include features identified in Section 4, and more importantly would leverage the interaction between these features to aid collaborative sensemaking. These features are: *visualizations: system generated, and user-generated; and information sharing: implicitly by the system, and user-mediated*. Further, based on the *design goals* identified in Section 5, the proposed solution would suggest recommendations to non-experts to aid sensemaking.

7 Expected Contributions

This doctoral work will have design contribution of creating an interface where micro-tasks may be auto-generated for the crowds and results produced are fed back into the original system for analysis. Theoretical contribution of my work would be understanding how micro-task based crowd-supported system may enhance task-resolution of complex tasks, and support collaboration.

References

[1] Bernstein, M. S. (2010, October). Crowd-powered interfaces. In Adjunct proceedings of UIST 2010 (pp. 347-350). ACM (2010).
[2] Bernstein, M.S., Little, G., Miller, R.C., et al. Soylent: a word processor with a crowd inside. Proceedings of UIST 2010, ACM(2010), 313–322.
[3] Chilton, L. B., Little, G., Edge, D., Weld, D. S., & Landay, J. A. (2013, April). Cascade: crowdsourcing taxonomy creation. In Proceedings of the CHI 2013 (pp. 1999-2008). ACM.
[4] Zhang, H., Law, E., Miller, R., Gajos, K., Parkes, D., & Horvitz, E. (2012, May). Human computation tasks with global constraints. In Proceedings of the CHI 2012 (pp. 217-226). ACM.

[5] Kittur,A., Smus,B., Khamkar,S., & Kraut, R.E. CrowdForge:Crowdsourcing Complex Work. Proceedings of UIST 2011, ACM(2011), 43.52

[6] Malone, T., Laubacher, R., & Dellarocas, C. (2009). Harnessing crowds: Mapping the genome of collective intelligence.

[7] Little, G., Chilton, L.B., Goldman, M., and Miller, R.C. TurKit: human computation algorithms on mechanical turk. Proceedings of UIST 2010, ACM (2010), 57–66.

[8] Kulkarni, A., Can, M., Hartmann, B. Collaboratively Crowdsourcing Workflows with Turkomatic.

[9] Lasecki, W. S., Wesley, R., Nichols, J., Kulkarni, A., Allen, J. F., & Bigham, J. P. (2013, October). Chorus: a crowd-powered conversational assistant. In Proceedings of the UIST 2013 (pp. 151-162). ACM.

[10] Goyal, N., Leshed, G., & Fussell, S. R. (2013). Effects of Visualization and Note-Taking on Sensemaking and Analysis. In Proceedings of the CHI 2013, ACM.

[11] Goyal, N., Leshed, G., Cosley, D., & Fussell, S. R. (2014). Effects of Implicit Sharing in Collaborative Sensemaking. In Proceedings of the CHI 2014, ACM.

[12] Goyal, N., Leshed, G., & Fussell, S. R. (2013). Leveraging partner's insights for distributed collaborative sensemaking. CSCW Companion 2013. ACM.

[13] Hu, Mengdie, et al. "OpinionBlocks: a crowd-powered, self-improving interactive visual analytic system for understanding opinion text." *Human-Computer Interaction–INTERACT 2013*. Springer Berlin Heidelberg, 2013. 116-134.

Understanding Collaborative Activities: A Distributed Cognition Perspective

Christina Vasiliou
Cyprus Interaction Lab
Dept. of Multimedia and Graphic Arts, Cyprus University of Technology
30 Archbishop Kyprianou Str., 3036 Lemesos
christina@cyprusinteractionlab.com

Abstract. Using a Distributed Cognition framework, this work aims to extend our understanding of how groups of learners collaborate in a learning environment rich in technologies.

Keywords. Human-Computer Interaction, CSCW, CSCL, Distributed Cognition, DiCoT

1 Introduction

Technology's evolving nature has brought new possibilities to the design of technology-rich learning environments for collaborative activities. Tablets and smartphones together with personal computers become part of a device ecology in which each device acts as a specialized portal into users' personal or shared information space [3]. In a collaborative learning environment these devices may be used for a variety of tasks while each individual may perform a task differently. Therefore, learners, tasks and use of technology in a learning environment cannot be studied independently or in isolation. To design effective technology-rich collaborative learning environments we need to acquire a deep understanding of the complex relations and interactions between collaborators and information technologies. Our work focuses on exactly this, using Distributed Cognition (DC) as a theoretical framework to guide data collection, analysis and interpretation of findings. Researchers have identified DC as a powerful tool for understanding the interdependencies of learners, tasks, and technologies in collaborative environments and for highlighting aspects for redesign [5]. The work is supervised by Dr. Andri Ioannou.

2 Theoretical framework

As a theoretical framework DC analyses cognition in a distributed manner [4], while it considers an activity taking place across individuals, tools and representations as one cognitive system, instead within an individual's mind.

Researchers in HCI (Human-Computer Interaction) and CSCW (Computer Supported Cooperative Work) communities have identified DC as a valid framework for understanding the interdependencies between users, tools and tasks [2]. In order to design effective technology rich learning environments, each tool must allow the distribution of an individual's cognition to the wider cognitive system, e.g. the classroom [7]. Furthermore, DC can provide a detailed identification of issues with existing work practices and mediating artefacts [6], allowing researchers to identify aspects of re-design of the environment.

However, there is no established methodology towards applying DC in the case of collaborative learning environments. In this work, we adopted the Distributed Cognition of Teamwork (DiCoT) methodology introduced by Blandford and Furniss [1] for collaborative work to extend our understanding of how groups of learners collaborate in a learning environment rich in technologies. This methodology draws together ideas from DC and conceptual design, including 22 principles classified loosely in three models; physical layout, information flow, and artefacts.

3 Proposed Research

Through the lens of DC, this work aims to extend our understanding of how groups of learners collaborate on design problems in a learning environment rich in physical and digital technologies and blending face-to-face content-and-activity with digital content-and-activity. Further, we aim to highlighting what is salient in the existing collaborative learning environment design and practices and indicate aspects of redesign of the learning environment.

For the purpose of this work, we enriched a postgraduate HCI course with four identical technology rich settings that aimed to support student collaborative activities around a design problem. Following an ethnographic approach, both qualitative and quantitative data were collected in active HCI courses over three consecutive years resulting in a rich dataset for analysis. In the sections that follow the progress achieved so far as well as the future plans for the current dissertation are outlined.

3.1 Work to date

3.1.1 Study 1: Understanding attitudes and technology use in a technology rich learning environment

This study included a pilot investigation of the technology rich learning environment, situated in a post-graduate HCI course. In order to understand learners' attitudes and technology use, we administrated questionnaires assessing students' motivational beliefs and overall satisfaction with the learning experience. Our results show that students' overall satisfaction was highly rated, while the technology rich learning environment contributed to students' engagement and collaboration [8].

3.1.2 Study 2: Understanding users' flow experience in a technology rich learning environment

This study focuses on the relationship between flow experience and the technology rich learning environment. The purpose was to explore learners' flow experience and to understand the affordances that engage users in the collaborative activities. Findings suggest that flow experience does exist in collaborative activities within the learning space, revealing individual affordances of the collection of devices used in the learning environment [9].

3.1.3 Study 3: Validating distributed cognition as a suitable framework for understanding collaborative learning activities in a technology rich environment

Using distributed cognition framework this study analyses learners' behaviour and how the technology rich learning environment supports collaboration and cooperation. The analysis allowed an in-depth understanding of the interactions among learners and tools during collaborative activities. Furthermore, the study validates distributed cognition as a suitable framework for understanding collaborative learning activities in a technology rich environment [10].

3.2 Future Research

3.2.1 Study 4: Constructing group profiles through the lenses of distributed cognition

The study will focus on constructing the individual profiles for each group validated distributed cognition as a modelling tool for collaborative learning activities through observation and video analysis.

3.2.2 Study 5: Distributed cognition in technology rich learning environments: Expanding the framework

Drawing from all the results of the previous studies, this study will expand the distributed cognition framework, including aspects referring to users' experiences and collaborative learning.

4 Expected Contributions

The completion of this project is expected to supply researchers a better understanding of how groups of learners collaborate in a learning environment rich in technologies. Furthermore, we aim to enrich and guide technology designers and practitioners on constructing effective technology rich environments for collaborative learning activities.

References

[1] Blandford, A., and Furniss, D. DiCoT: A methodology for applying distributed cognition to the design of team working systems. In: Gilroy, S.W., Harrison, M.D. (eds.), DSV-IS 2005. LNCS, vol. 3941, pp. 26–38. Heidelberg: Springer, 2006.

[2] Halverson, C.A. Activity theory and distributed cognition: Or what does CSCW need to Do with theories? Computer Supported Cooperative Work 11, 243–267. 2002.

[3] Houben, S., Tell, P., and Bardram, J. E. ActivitySpace: Managing Device Ecologies in an Activity-Centric Configuration Space. In Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces, pp. 119-128. ACM, 2014.

[4] Hutchins, E. Cognition in the wild. Cambridge: MIT Press, 1995.

[5] Loke, S. W., and Ling, S. Analyzing Observable Behaviours of Device Ecology Workflows. In ICEIS (4), pp. 78-83, 2004.

[6] Rogers, Y. HCI Theory: Classical, Modern, and Contemporary. Synthesis lectures on human-centered informatics. Morgan & Claypool, San Francisco, 2012.

[7] Valanides, N., and Angeli, C. Distributed cognition in a sixth-grade classroom: An attempt to overcome alternative conceptions about light and color. Journal of Research on Technology in Education 40 (3), pp. 309–336. 2008.

[8] Vasiliou, C., Ioannou, A., and Zaphiris, P. Technology Enhanced PBL in HCI Education: A Case Study. In Human-Computer Interaction–INTERACT 2013, pp. 643-650. Springer Berlin Heidelberg, 2013.

[9] Vasiliou, C., Ioannou, A., and Zaphiris, P. Measuring Students' Flow Experience in a Multimodal Learning Environment: A Case Study. In Learning and Collaboration Technologies. Designing and Developing Novel Learning Experiences, pp. 346-357. Springer International Publishing, 2014.

[10] Vasiliou, C., Ioannou, A., and Zaphiris, P. Understanding collaborative learning activities in an information ecology: A distributed cognition account. Computers in Human Behavior, 41, pp. 544-553. 2014.

Accessibility for Cognitive and Mentally Disabled Persons

Stefan Johansson
Royal Institute of Technology, CSC – School of Computer Science and Communication, Sweden
Stefan.johansson@funka.com

Abstract. This document provides a short description of my current work on how persons with mental and cognitive disabilities cope in the digital society.

1 The research area or sub-area of the work

Accessibility for cognitive and mentally disabled persons

1.1 A brief description of the research topic

Is the society digitally accessible for persons with mental disabilities? It is likely that the fast development of devices, services and techniques open up opportunities for a better life but it is also likely that different kind of obstacles may occur. Is there a digital divide between citizens in general and citizens with mental disabilities? And if so; what is the nature of this divide?

How can persons with mental disabilities take part in research studies and how can they take part in processes in order to change or produce new services or devices?

For further information about the topic, please see my paper 84 - Cognitive accessibility for mentally disabled persons, which has been accepted for publication and presentation at INTERACT 2015.

2 Description of the research problem

The problem is of two kinds. The first is to explore and describe how persons with mental illnesses could participate and benefit from possibilities created by what we in a broad term can call the digital society. The second is how to do this research in itself. Traditional methods for user participation needs to be adjusted end developed when users suffer from different kinds of mental health problems.

2.1 The research hypothesis

There is a gap between citizens in general and citizens with mental illness in terms of being involved in the digital society.

Persons with mental illnesses are often excluded from research in the field of HCI. This is especially true when it comes to main stream subjects.

2.2 Methods

The overall research approach follows the tradition of Action Research. More specifically the method is heavily influenced by the tradition of Participatory Design or the more Scandinavian version of Cooperative Design and we have therefore used an adapted form of Participatory Action Research.

In the study described in paper 84 - Cognitive accessibility for mentally disabled persons, we have used a concept called study circles. Study circles have a long tradition in the Nordic countries. A study circle is a group of adults that meet and discuss a specific topic. It is done by free will and in a democratic way. There is no teacher; instead a study circle has a leader who facilitates the discussions. The concept of study circles has been complemented with methods to visualize key findings and key conclusions.

In another study we work with persons that also have cognitive disabilities but in this cased caused by adhd, dyslexia, autism and intellectual disabilities. In the study we test different kind of data collection methods to evaluate how easy it is for the participants to take part.

3 A sketch of the proposed solution

The solution I am proposing is to work with a mix of data collection methods. I have started to create cognitive user test profiles based on how individuals say they can take part in research and development processes. So far we have tested 20 different data collection methods. By offering 3-5 methods it seems that every participant can pick at least one method that makes it possible for them to participate. This indicates that the design of research projects could benefit from offering different ways of participation.

3.1 The expected contributions of the PhD research

I expect taking part in the DC will help me to be more precise and to focus on the right thing. Having the opportunity to present and discuss my research and doing the same with others will help me to sharpen my arguments and to present my material with more rigor end relevance.

3.2 A statement of work to date and open questions/issues for discussion at the DC.

I would like to discuss methods of user participation when users have cognitive and mental disabilities. Those users can be difficult to involve both because of their problems but also because of the research community having difficulties in adapting methods that do work. A person with impairment cannot do much about the fact that there is an impairment but the research community should be able to do a lot to develop new methods or adapting old methods so users can be involved.

I would also like to discuss how HCI research in general could take cognitive accessibility into account. Systems, interfaces and devices are often too difficult to use and the potential for simplicity that lies hidden in a cleverly designed product or service is often lost.

Work to date is that one study is finished and will be presented at the Interact 2015. Another study is in year 2 of a 3 year project, where the work with testing data collection methods will end at 2 year mark (in august 2015). A third project where we will test methods for user participation together with homeless persons and the social service in Stockholm, Sweden has just started.

Multimodal Approaches for Text Entry in Indian Language on Mobile Devices

Sanjay Ghosh
Indian Institute of Technology, Bombay,
Powai, Mumbai, India
sanjayghosh@iitb.ac.in

Abstract. Text input process in Indic scripts is complex due to the large set of complex characters and more number of user actions required to enter a syllable. Research objective is to identify and evaluate the effective ways of text input in Indic scripts on mobile devices through usage of multimodal interactions. Scope of this research includes identifying various problems faced by the users in entering Indic language text on touchscreen mobile devices, theoretically analysing few of these problems, and designing solutions for some of those through appropriate use of multimodal interactions. The overarching objective of this proposed research work is to contribute towards making the text entry in Indic script on mobile devices easy and faster. This doctoral thesis is under the supervision of Prof. Anirudha Joshi and is a part of the on-going research on Indian language text entry at IIT Bombay, India.

1 Introduction to the Research Topic

In India, with around 75% average literacy and very less English language writing competency, a better solution for text entry in several Indic languages on touch screen phones is the need of hour. This has become a crucial area of HCI research. Typing of Indic scripts or for that matter any Alphasyllabary script involves composing of a syllable by entering consonant and vowel modifiers. Additionally there are a large set of complex characters such as conjuncts, diacritic marks, etc. Thus, text input process in these scripts requires more number of user actions. Now, in a typical multimodal interaction, individual modalities play their well-coordinated roles, making them appropriate and natural for different types of tasks. Therefore, it needs to be investigated whether it will be more effective if these different actions be performed using different modalities combination. In this on-going study the following research questions were under exploration –

RQ1: Does the simultaneous use of more than one input interaction modalities improve the effectiveness of text entry in Indic script on a mobile device?

RQ2: During text entry in Indic script on a mobile device, what are the effective roles of different input modalities?

- Prior literature [1], [2] mentions several issues with typing on Indic language keyboards such as, complexity of the Indic script, longer character search time, lower accuracy, involving higher cognitive load and most importantly, very slow rate of input. Work by Joshi et al. [3], and Lauren et al. [4] suggested different approaches related to the soft keyboard layouts to improve text entry in Indic script. In terms of the problem space, our exploration fairly aligns to the aforementioned researches. Classical works by Zhai et al. [5] and MacKenzie [6] have provided theoretical modelling of few text entry methods on touch screen keyboards for English script. Now, one of the challenges in applying these modelling techniques for Indic language scripts lie in the way Indic scripts input varies from English script input. In the scope of the current research we would also investigate methods to model the text entry for Indic scripts, similar to the classical works [5], [6] on touch screen keyboards for English script.

2 Methodology and Hypothesis

As this research problem involves dual aspects, i.e. exploring and identifying the problem space (inductive) as well as defining the solution space and evaluating solutions (deductive), it calls for employing multi method studies or Mixed Method. The work till date involved only the inductive approach wherein we performed field study with 50 users from the target user group to identify and categorize various issues as well as to observe their typical usage pattern while entering Indic language text on mobile. Additionally, in the inductive exploratory phase, we performed few controlled experiments with 18 participants to understand their usage patterns while performing few tasks using multimodal interactions. At this juncture, from our field study observations we formulated the following operational hypothesis for this research.

- H1: Simultaneous use of multimodal input mechanism improves the rate of text entry in Indic language.
- H2: Effective use of multiple input mechanisms for text entry may be realized when each of the modalities perform different task, such as being used for composition of different parts of a syllable

Going forward we plan to theoretically model the process of Indic language text entry on mobile device. This modelling exercise would suggest which are the steps involved in Indic language text entry that can potentially be

improved using multimodal interaction. For few of those scenarios that theoretically indicate potential improvement, we'll design prototype solutions. Following that in the deductive approach, we plan to conduct controlled experiments with the target users through the use of those prototypes.

3 Research Progress

We started our exploration by defining the preferred user group for our study. This involved people who could write in any preferred Indian language better than they could do so in English along with moderate to high exposure of mobile phones. We then performed an open ended field study with 50 participants from the preferred user group with an objective to identify and categorize the key issues that users face in using Indian language keyboard [7]. Four totally different kinds of virtual mobile keyboards were used as stimuli in these sessions. An exhaustive list of 46 issues and points observed during the study as well as those reported in prior research were categorized into 4 categories; those due to the complexity of the script, keyboard interface, user's cognitive performance and social aspects.

On a parallel track we performed few controlled experiments wherein the users were asked to perform few tasks such as navigation and text editing by making simultaneous use of multiple interaction modalities. We evaluated their level of performance, accuracy and user experience [8] as well as analysed few of their usage patterns such as patterns of errors, modality switching and modality preferences [9]. We analysed the steps required to input various types of characters namely, consonants, matras (vowel marks), conjuncts, diacritic marks for four popular Indic language keyboards.

Our current activity involves theoretical modelling and analysing maximum typing performance on various popular Indian language soft keyboards on mobile. Challenge lies in extending and applying the theory of movement time based on the Fitts' Law and reaction time based on the Hick-Hyman's Law reported for English keyboard onto Indic language text entry methods.

4 Future work and expected contributions

Currently our research is in the stage of refining our preliminary hypothesis through theoretical modelling and exploration of few prototype solutions. Prior to making effort in directly creating various multimodal text entry prototypes, we are theoretically evaluating potential solutions. We would investigate on performance of multimodal input keyboards by first designing

few multimodal methods of text input on hypothetical keyboards, theoretically analysing the performance on those finally experimentally evaluating the same.

We believe that our research can offer the following contributions in terms of new knowledge on: 1) methods to improve effectiveness of text input; 2) solving few issues encountered during Indic language text entry; 3) theoretical modelling of text entry method for Indic script; 4) interplay of multimodal input interactions for text input.

5 Open Questions to be discussed

Here we mention few open questions and discussion points for which we would like to seek some suggestions from the community of experts.

- How to extend the theoretical modelling of text entry in Indic language?
- How much is the validity of the preliminary hypothesis mentioned earlier?
- How much is the practical viability of text entry using multimodal?
- Comment on the scope of this research.

References

[1] Joshi, A., Ganu, A., Chand, A., Parmar, V., and Mathur, G. Keylekh: a keyboard for text entry in indic scripts. In Proceedings of CHI extended abstracts on Human factors in computing systems, ACM, 2004, pp.928-942.

[2] Samanta, D., Sarcar, S., and Ghosh, S. An Approach to Design Virtual Keyboards for Text Composition in Indian Languages. International Journal of Human-Computer Interaction, 29(8), pp. 516-540.

[3] Joshi, A., Dalvi, G., Joshi, M., Rashinkar, P., and Sarangdhar, A. Design and evaluation of Devanagari virtual keyboards for touch screen mobile phones. In Proceedings of 13th International Conference on Human Computer Interaction with Mobile and Services, ACM, 2011, pp. 323-332.

[4] Hinkle, L., Brouillette, A., Jayakar, S., Gathings, L., and Kalita, J. Design and Evaluation of Soft Keyboards for Brahmic Scripts. Transactions on Asian Language Information Processing, 12(2), ACM, 2013.

[5] Zhai, S., Hunter, M., and Smith, B. A. The metropolis keyboard-an exploration of quantitative techniques for virtual keyboard design. In Proceedings of the 13th annual ACM symposium on User interface software and technology, ACM, 2000, pp. 119-128.

[6] MacKenzie, I. S. A note on the information-theoretic basis for Fitts' law. Journal of motor behavior, 21(3), 1989, pp. 323-330.

[7] Ghosh, S., and Joshi, A. Text Entry in Indian Languages on Mobile: User Perspectives. In Proceedings of the India HCI 2014 Conference on Human Computer Interaction ACM, 2014, pp. 55.

[8] Ghosh, S., Joshi, A., and Tripathi, S. Empirical evaluation of multimodal input interactions. Human Interface and Management of Information, Springer Berlin Heidelberg, 2013, pp. 37-47.

[9] Ghosh, S., and Joshi, A. Exploration of Usage Patterns for Multimodal Input Interactions. In Proceedings of the India HCI 2014 Conference on Human Computer Interaction ACM, 2014, pp. 64.

A Multi-Modal System for Public Speaking

[1]Fiona Dermody [2]Dr. Alistair Sutherland [3]Dr. Margaret Farren
[2,3](Supervisors)
[1,3]Faculty of Humanities
[1,2]Faculty of Engineering and Computing
Dublin City University, Dublin 9. Ireland.

Fiona.Dermody3@mail.dcu.ie

Abstract. This research involves the development of a digital system and user interface to analyze social signals displayed during public speaking and to provide real-time feedback to users on speaking performance.

Keywords: Human Computer Interaction, Affective Computing, Multi-Modal Interfaces

1 Research Topic Details

1.1 Description of Research Problem to be solved

Can a digital system be developed to recognize the multi-modal social signals of effective public speaking and through the provision of feedback enable users to become more effective at public speaking?

1.2 Why it is important to solve this research problem

The fear of public speaking tops the list of human phobias. However, success in social, academic and occupational situations depends on the ability to communicate effectively to groups. A fear of public speaking thus limits achievement in social gatherings, education and enterprise [1]. However, the fear of public speaking can be so great that it can lead to avoidance of speaking in the public domain altogether [2]. The problem is a recursive one, how can an individual improve their speaking skills and reduce their fear of speaking in public if they avoid speaking in public?

Public speaking is not just about the words spoken. Effective speaking involves the use of gestures, facial expressions and vocal variety. All these

social signals combine to give the appearance of self-confidence in a speaker. Research has found that anxious speakers do not engage the attention of an audience [3]. Rather the audience focuses on the speaker's nervous disposition instead of their words [4].

The solution, which we propose, is interdisciplinary. It incorporates theory from computing, psychology and communications. It is envisaged that this social signal recognition system will enable individuals to develop their competence in public speaking.

Using a combination of 3D video imaging, audio and social signal processing algorithms, this digital system analyses facial expressions, tone of voice and gestures. The system then provides feedback on the user's speaking performance. It will also deliver tutorial videos on good speaking practices. Exposure to these dynamic features will enable a speaker to systematically develop confidence and skill before speaking in front of a live audience.

One of the technical challenges to be overcome during the development of the system is the optimal way to display feedback to the user in real time during their speaking task. The nature of the feedback is imperative, as our survey showed that some users prefer visual feedback while others prefer textual feedback. We decided to allow feedback to be customizable on the interface.

Skilled human trainers in communication are scarce and expensive. This digital system incorporates experience from one such skilled human trainer to provide constructive feedback to users on their speaking performance. This digital experience will enable anxious speakers to develop their public speaking skills cost-effectively, in private and at their own pace.

By harnessing the power of social signal processing, this system will increase the user's communication skills, confidence and ultimately lead to greater success in life.

1.3 Justification that prior research has not solved this problem

There have been attempts to use social signal recognition for public speaking but not for instruction purposes [5]. This system is innovative because it will extend the field of human computer interaction:

- Combining all modalities - voice, gesture, facial expression and body pose into a multi-modal system for delivering instruction in public speaking
- Incorporating the knowledge of experts in public speaking ensures that the feedback provided on a user's speaking is based on a real-world, practice-based approach

2 Research Hypotheses

- A multi-modal system can accurately recognise the characteristics of gestures, body language, voice, facial expression that influence an audience during public speaking
- Real-time multi-modal feedback is most effective for users to develop skill in public speaking

3 Methods used to develop and evaluate system

3.1 Prototype Development

A prototype has been developed following a user survey on the features required in a multimodal system for public speaking.

3.2 Evaluating the work and presenting credible evidence of results to the research community

- Using the knowledge of experts in public speaking to construct a tutorial system, which will give automatic feedback
- Carrying out controlled experiments on volunteer users to measure the efficacy of the system.

4 Sketch of the proposed solution

The system will use a Microsoft Kinect connected to a computer. Social Signal Processing techniques will be used to recognize the speaker's body language, gestures, voice, and facial expressions

- Classify speaker's emotion, as perceived by the audience, from the combination of the above
- Perceived emotion is regarded as the primary component for analysis.
- Give feedback to user on speaking performance
- Provide examples of good and bad speaking practice
- Set tutorial exercises from beginners level to advanced level and will evaluate the user's performance

5 Expected contributions of the PhD Research

There have been attempts to use social signal recognition for public speaking but not for instruction [5]. Our research will extend the field of human computer interaction -

- To combine all modalities (voice, gesture, face, etc.) into a multi-modal system for public speaking
- To construct models for different types of speech, speaking styles, contexts
- To use the knowledge of experts in public speaking to construct a tutorial system, which will give automatic feedback
- To carry out controlled experiments on volunteer users to measure the efficacy of the system.

6 Statement of work to date

- User survey on features required in multimodal system for public speaking
- A prototype has been developed
- Initial user testing has started

7 Issues for discussion at the Doctoral Consortium

- Prototype interface will be presented for discussion
- Is prototype interface design effective?
- Is the feedback presented in a clear and effective way?

Acknowledgements

This material is based upon works supported by Dublin City University under the Daniel O'Hare Research Scholarship scheme.

References

[1] Dwyer, K. K., & Davidson, M. M. Is Public Speaking Really More Feared Than Death? Communication Research Reports, 29(2), 2012, pp. 99-107.

[2] Feldman, P. J., Cohen, S., Hamrick, N., & Lepore, S. J. Psychological stress, appraisal, emotion and cardiovascular response in a public speaking task. Psychology & Health, 19(3), 2004, pp. 353-368.

[3] Freeman, T., Sawyer, C. R., & Behnke, R. R. Behavioral inhibition and the attribution of public speaking state anxiety. Communication Education, 46(3), 1997, pp. 175-187.

[4] Winton, E. C., Clark, D. M., & Edelmann, R. J. Social anxiety, fear of negative evaluation and the detection of negative emotion in others. Behaviour Research and Therapy, 33(2), 1995, pp. 193-196.

[5] Batrinca, L., Stratou, G., Shapiro, A., Morency, L. P., & Scherer, S. Cicero-towards a multimodal virtual audience platform for public speaking training. In Intelligent Virtual Agents, Springer Berlin Heidelberg, 2013, pp. 116-128.

Improving Accessibility Support for Web Developers[1]

David Swallow
University of York
Human Computer Interaction Research Group, Department of Computer Science,
University of York, York, YO10 5GH, United Kingdom
david.swallow@york.ac.uk

Abstract. Despite many initiatives over the last decade to improve web accessibility for people with disabilities, the accessibility of websites has barely improved. This PhD research examines web developers' mental models to understand why they are failing to create accessible websites and what can be done to better support them. It describes the development and evaluation of a new accessibility information resource for assisting web developers in the creation of accessible websites and applications.

Research area. Web accessibility, web development, and mental models

1 Research problem and hypotheses

Web developers have an obligation to develop websites that are accessible and usable by the broadest range of users, including people with disabilities. Over the last decade, there have been many initiatives to improve the accessibility of websites for people with disabilities. This has resulted in a well-established body of accessibility information, often presented in the form of a set of guidelines or recommendations. Despite these initiatives however, the accessibility of websites has barely improved during this period and, according to certain studies (e.g. [4], [6]), has worsened.

The second version of the Web Content Accessibility Guidelines, WCAG 2.0 [8], comprises a complex suite of documents that assumes the reader understands a considerable amount about people with disabilities and the assistive technologies they use. There are numerous studies (e.g. [1-3]) and much anecdotal evidence suggesting that web developers are extremely confused by the guidelines. What is less evident is what is the nature of the confusion and why that is resulting in inaccessible websites.

---

[1] This dissertation work is co-supervised by Prof. Helen Petrie (helen.petrie@york.ac.uk) and Dr. Christopher Power (christopher.power@york.ac.uk) at the University of York, UK.

One approach may be to focus upon web developers' mental models of the web development process and how the needs of disabled web users can be fitted into these. A study by LaToza, Venolia and DeLine [5] concluded that software developers go to great lengths to create and maintain rich mental models of code that they rarely explicitly record. Web developers undoubtedly have mental models of their web development process, reflecting their knowledge and understanding of the process. However, it is unclear whether their mental models incorporate an adequate understanding of disabled users and how best to support them through accessible coding practices. By examining web developers' current mental models, I will understand not only their knowledge of web development but also their awareness of web accessibility and the needs of disabled users.

My hypotheses are:

- Exploring web developers' understanding of web accessibility and web development practices will provide an indication of the misconceptions, errors and knowledge gaps that result in inaccessible websites.
- Developing an accessibility information resource that addresses the misconceptions, corrects the errors and fills in the knowledge gaps will support web developers to form more useful and maintainable mental models and foster a greater understanding of web accessibility.

2 Proposed method and solution

The first stage of my research explores the current state of web accessibility and the existing working practices of professional web developers. It also establishes their needs and requirements with regards to web accessibility resources, and gains an understanding of their mental models of web development processes and web accessibility. Three key findings emerged from this stage as to why web developers struggle with accessible web development. The first is that existing accessibility information resources use domain-specific terminology with which web developers are often unfamiliar. The second is that the organisation of existing resources is often different to how web developers approach web development. The third is that existing resources tend to overwhelm web developers by presenting too much information at once.

The second stage of my research focuses upon what can be done to improve web developers' understanding of web accessibility. It documents the development of an accessibility information resource, called WebAIR (Web Accessibility Information Resource) that I have designed to change and refine the mental models of web developers and foster a greater understanding of

disabled users and web accessibility. The design of WebAIR corresponds to the findings from the first stage of my research. WebAIR avoids domain-specific web accessibility terminology and instead refers to web development terms or specific user actions in the interface. WebAIR accommodates web developers' existing approaches to web development, specifically their tendency to structure their work according to the types of web content on which they are working. WebAIR avoids overwhelming web developers with information by presenting a limited selection of example solutions drawn from the many techniques provided in WCAG 2.0 [8]. WebAIR also provides just-in-time training in web accessibility concepts, allowing web developers to gradually learn about the domain of web accessibility and why they are undertaking specific web accessibility tests.

The third stage of my research evaluates the effectiveness of WebAIR in changing and refining web developers' mental models of web accessibility and web development practices. This stage also explores whether the resource can be integrated into the web development process to promote the long-term maintenance of useful mental models.

3 Work to date

The work I have completed to date includes:

- A contextual inquiry investigation with 13 professional web developers, and an online survey, completed by 60 respondents from across Europe. The aim of this was to explore the role of web developers in much greater detail in order to understand their working practices and establish a set of requirements for providing accessibility support.
- An interview-based study of web developers' mental models, conducted with 26 professional web developers. The aim of this was to explore their understanding of disabled users, web accessibility and web development practices.
- An initial validation of my proposed solution, WebAIR, with 26 professional web developers and 7 student web developers [7]. The aim of this was to validate WebAIR with both professional and student web developers and explore their initial impressions, using a rating scale and a short interview.
- An evaluation of WebAIR with 50 student web developers. The aim of this was to determine the effectiveness of WebAIR in allowing student web developers to identify and address accessibility problems.

I am currently planning a further evaluation of WebAIR with professional web developers. The aim of this is to explore how they use the resource and

whether it corresponds to their mental models of web accessibility. Participants in this study will be recruited to take part in a longitudinal diary study of WebAIR, which will further explore the effectiveness of WebAIR and investigate the long-term maintenance of useful mental models.

4 Contributions and open questions

The expected contributions of my research include: an exploration of the mental models and working practices of web developers; greater insight into why web developers are failing to create accessible websites; an understanding of the accessibility information necessary to change and refine web developers' mental models; and an evaluation of the effectiveness of WebAIR in fostering a greater awareness and understanding of web accessibility.

At the Doctoral Consortium, I would like to share my experiences of conducting studies "in the wild" with professional web developers. I would also like to discuss the value of examining mental models of web accessibility and the robustness of the methods I have used to elicit, analyse and interpret them.

References

[1] Alonso F, Fuertes J, González Á, and Martínez L. Evaluating Conformance to WCAG 2.0: Open Challenges. In: Miesenberger K, Klaus J, Zagler W, Karshmer A, eds. Computers Helping People with Special Needs. Springer Berlin / Heidelberg, 2010. pp. 417–24.

[2] Alonso F, Fuertes J, González Á, and Martínez L. On the testability of WCAG 2.0 for beginners. Proceedings of the 2010 International Cross Disciplinary Conference on Web Accessibility (W4A). Raleigh, North Carolina: ACM, 2010. pp. 1–9.

[3] Brajnik G. Validity and reliability of web accessibility guidelines. Proceedings of the 11th international ACM SIGACCESS conference on Computers and accessibility. Pittsburgh, Pennsylvania, USA: ACM, 2009. pp. 131–8.

[4] Hackett S, Parmanto B, and Zeng X. A retrospective look at website accessibility over time. Behaviour and Information Technology. 2005, 24(6), pp.407–17.

[5] LaToza TD, Venolia G, and DeLine R. Maintaining mental models: a study of developer work habits. Proceedings of the 28th international conference on Software engineering. Shanghai, China: ACM, 2006. p. 492–501.

[6] Loiacono ET, Romano Jr. NC, and McCoy S. The state of corporate website accessibility. Communications of the ACM. 2009, 52(9), pp. 128–32.
[7] Swallow, D, Power, C, Petrie, H, Bramwell-Dicks, A, Buykx, L, Velasco, CA, Parr, A and O Connor, J. Speaking the Language of Web Developers: Evaluation of a Web Accessibility Information Resource (WebAIR). In Miesenberger K, Fels D, Archambault D, Penaz P, and Zagler W, eds. Computers Helping People with Special Needs. Springer Berlin / Heidelberg, 2014. pp. 348-355.
[8] World Wide Web Consortium. Web Content Accessibility Guidelines (WCAG) 2.0. 2008. http://www.w3.org/TR/WCAG20/

Towards a Methodology to Evaluate Multimodal Games for Cognition in People who are Blind

Ticianne Darin
Federal University of Ceará
Av. Humberto Monte S/N 60455900 Fortaleza, Ceará, Brazil

Abstract. Multimodal serious video games are relevant tools to enhance orientation and navigation skills in people who are blind. We intend to propose a methodology for evaluating the usability and the cognitive impact of interface elements and interaction components in this kind of games. The results will allow us to analyse and discuss the impact of interface elements on cognition and propose guidelines for the better use of these elements.

1. Introduction

1.1 Research Problem

People who are blind face difficulties to complete tamethosks requiring spatial representation [2]. As a result, they tend to walk choosing the safer route rather the most efficient one [4,9]. In order to navigate efficiently, this audience needs non-visual stimuli to perceive the environment and construct effective cognitive maps [3]. Multimodal interfaces can stimulate orientation and navigation skills in this population by using their complementary channels via audio-based interfaces or haptic/kinesthetic feedback [10, 11, 14]. Studies have shown that multimodal video games can stimulate cognitive processes such as tempo-spatial orientation and haptic perception [13,16]. Since children and young people widely use games as part of their daily routine [5], multimodal serious games are attractive tools to stimulate cognitive development and improvement.

A previous study [12] analyses the state of the art in design, evaluation and the existing technologies for multimodal video games, designed to enhance cognition of people who are blind. They state that it is necessary to identify the advantages and limitations of the evaluation methods for multimodal interfaces in this context. Regarding usability, the results show that there is no consensus about the elements to evaluate, nor the methodology and evaluation instruments. Besides, it is fundamental to ensure that these games can stimulate the cognitive development, but cognitive impact evaluations are still

scarce. Designers and developers need to understand how to use multimodal elements in a relevant and meaningful way, to create and to evaluate a serious game designed to enhance cognitive skills. In order to meet this need, the main objective of this work is to propose a methodology for evaluating interface elements and interaction components in multimodal serious games designed to stimulate cognitive skills in people who are blind.

1.2 Research Hypothesis

- We assume that it is possible to design a model for evaluating the usability and the cognitive impact of a serious multimodal video game, combining qualitative, quantitative and quali-quantitative approaches. The main goal of this research is to propose an appropriate and efficient methodology that suits these requirements.

The existence of such a methodology implies there are specific interface and interaction elements that are significant to the construction of cognitive maps, and thus cognition in people who are blind. Determining which are these elements will allow us to describe the best practices for their adequate use.

2. Methodology

- The methodology of this research consists of six main phases, as follows:
  - **Literature review:** We already executed a Systematic Review [8] in order to identify the design and evaluation approaches, and the technologies currently in use for multimodal serious games, designed to support the development of mental maps, cognitive spatial structures and navigation skills.

Selection of Games: From the 21 applications studied in the Systematic Review, we will select a set of six multimodal video games to compose the testing sample (three casual and three serious). We will work with existing applications.

Selection of Evaluation Methods: We will define a set of quantitative, qualitative and quali-quantitative methods and instruments for usability and cognitive impact evaluation, such as Cognitive Walkthrough [17], Heuristic Evaluation [6], End-user and Facilitator Questionnaire for Software Usability [15] and quasi-experimental designs [1] for impact evaluation [3, 14].

Evaluation of the Methods: In order to identify the advantages and disadvantages of the methods to evaluate the usability and cognitive impact in

this particular context, we will evaluate each selected video game with diverse qualitative and quantitative methods.

Results Compilation: The evaluation results will be interpreted according to statistical analysis, as well as with the observations of the comparison groups and cross-site analyses. As a result, we will propose an appropriate and efficient methodology for the evaluation of usability and cognitive impact of multimodal video games for people who are blind.

Proposal Validation: We will use our methodology to evaluate the usability and the cognitive impact of the selected games. We will compare the results of our proposal with the results obtained with the evaluation methods, in order to determine the effectiveness of the proposed methodology.

3. Solution and Contributions

The proposed solution, as well as our major contribution, will be an appropriate and efficient methodology for the evaluation of usability and cognitive impact of multimodal video games designed for the enhancement of cognition in people who are blind. Besides, the nature of the research allows us to state the impact of user interfaces and multimodal interaction for the construction of cognitive maps in these learners, through interaction with multimodal games. Thus, as a final contribution, we expect to supply the designers and developers with guidelines for suitable interface and interaction choice for the design of multimodal video games, in the context of the cognitive development in learners who are blind.

4. Discussion

Some issues related to the proposed methodology should be considered, in order to improve the research findings. The first one is related to what are the best criteria for the selection of the evaluation methods. Another important point is how many and which methods would be the most relevant to evaluate. Finally, we would like to consider whether including both usability and cognitive impact evaluation methods in the same evaluation methodology is actually the best approach.

Acknowledgement

This paper was supervised by the professors Rossana Andrade and Jaime Sánchez and was funded by the Program STIC-AmSud-CAPES/CONYCIT/MAEE, project KIGB-Knowing and Interacting while Gaming for the Blind, 2014.

References

[1] Campbell, Donald Thomas, Julian C. Stanley, and Nathaniel Lees Gage. Experimental and Quasi-Experimental Designs for Research. Boston: Houghton Mifflin, 1963.

[2] Kolb, B and Whishaw, I. Neuropsicología Humana 5a edición. Editorial Médica Panamericana, 2006.

[3] Lahav, O. and Mioduser, D. Construction of Cognitive Maps of Unknown Spaces using a Multy-Sensory Virtual Environment for People who are Blind. Computers in Human Behavior 24(3), 2008, pp. 1139-1155.

[4] Lahav, O. and Mioduser, D. Multisensory Virtual Environment for Supporting Blind Persons' Acquisition of Spatial Cognitive Mapping – A Case Study. In World Conference on Educational Multimedia, Hypermedia and Telecommunications. 2001. pp. 1046-1051.

[5] Mayo, M. Games for Science and Engineering Education. Communications of the ACM 30(35), 2007, pp. 30–35.

[6] Nielsen, J. Heuristic Evaluation. Usability Inspection Methods. New York: John Wiley & Sons, 1994.

[7] Nikolakis, G. Tzovaras, D. and Strintzis, M. Object Recognition For The Blind. In Proceedingf of 13th European Signal Processing Conference (EUSIPCO2005), Antalaya, Turkey, 2005.

[8] Petersen, K., Feldt, R., Mujtaba, S. and Mattsson, M. Systematic mapping studies in software engineering. In Proceedings of the 12th Intl. Conference on Evaluation and Assessment in Software Engineering, Italy. 2008

[9] Pressl, B. and Wieser, M. A Computer-Based Navigation System Tailored to the Needs of Blind People. Lecture Notes in Computer Science, Volume 4061, 2006.

[10] Sánchez, J. User-Centered Technologies for Blind Children. Human Technology Journal, 45(2), November 2008, pp. 96-122

[11] Sánchez, J. and Tadres, A. Audio and Haptic Based Virtual Environments for Orientation and Mobility in People Who are Blind. In Proceedings of the 12th International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '10). ACM, NY, 2010.

[12] Sánchez, J., Darin, T. and Andrade, R. Multimodal Videogames for the Cognition of People who are Blind: Trends and Issues. In Proceedings of the 17th International Conference on Human-Computer Interaction (HCII'15), 2015. *(in press)*

[13] Sánchez, J. and Sáenz, M.: Usability of Audio-Based Virtual Environments for Users with Visual Disabilities. In Virtual Reality and Human Behavior Symposium, Laval, France, 2007.

[14] Sánchez, J. and Tadres, A. Audio and Haptic based Virtual Environments for Orientation and Mobility in People Who are Blind. In Proceedings of the 12th international ACM SIGACCESS conference on Computers and accessibility (ASSETS '10). ACM, NY, USA. 2010

[15] Sánchez, J. End-User and Facilitator Questionnaire for Software Usability. Usability evaluation test. University of Chile, 2003.
[16] Sjostrom, C. Using Haptics in Computer Interfaces for Blind People. In Proceeding of the ACM CHI 2001, Seattle, Washington, USA, March 31 – April 5, 2001 pp. 245–246
[17] Wharton, C., Rieman, J., Lewis, C,, and Poison, P. The Cognitive Walkthrough Method: A Practitioner's Guide. In Nielsen, J., and Mack, R. L. (Eds.), Usability Inspection Methods, John Wiley & Sons, New York, 105–140. 1994.

Bridging the Digital Gap in South Africa: Fitts's Model Impact on Usability

Guy Toko, Ernest Mnkandla
University of Johannesburg
University of South Africa
17 Bunting Road, AucklandPark Johannesburg, South Africa
gtoko@uj.ac.za

Abstract. Handheld computing systems are relatively new compared to existing computing systems. While their miniature size makes them ideal for mobility and connectivity, it also poses an enormous challenge to first-time users. For some, it may well be the solution to the huge digital gap that exists in South Africa. For others, the solution to the digital gap relates to issues of in computing usability, and in increasing the numbers of new or first-time users mostly in remote and impoverished communities. When it comes to systems usability, Fitts's law enables To assess the success rates, numbers of usability errors, the amount of frustration that any user may be experiencing., This model would be used to determine whether or not first-time users are able to overcome the usability challenges associated with handheld systems, and if in the way, the existing usability gap can be reduced.

1 Research area

Human computer interaction; First-time user; Adaptation; Remodelling Fitts's law;

1.1 Brief Description

This research project is about the impact and re-modelling of the Fitts's law on first-time handheld users in developing countries such as South Africa.

1.2 Description of the research

In the past few decades, new and innovative technologies in the field of ICT has emerged with specific regard to mobile handheld devices such as the tablet computing system (Tegarden & Dennis & Wixom, 2013). This appears to have reactivated the debate on systems usability, efficiency and effectiveness. Developers are not always aware of some of the challenges of first-time users, especially those who had never used any computing system in their lives (Lazar, 2007). Most, if not all of these devices, are imported from developed nations with different social, economic and technological realities. User needs

and aspirations are meant to be at the centre of usability studies for any computing systems development Indeed, it is expected that people everywhere in the world, including those in remote and impoverished places would have access to computing devices to enable them perform simple daily tasks (Lazar & Feng & Hochheiser, 2010). The question then is: if all potential users are provided handheld computing system in an attempt to close the digital gap, would they be able to such devices effectively?

Some scholars claim that the polarisation of societies makes it difficult if not impossible for every individual to have access to the same technological tools and facilities thus rendering such people technologically isolated (Reiss, 2012). The challenge is to narrow the existing technological gap by enabling access to new users at a higher rate thus increasing the number of first-time users who are in the process of mastering available ICT tools in a reduced time frame (Preece & Rogers & Sharp, 2002). For a number of years, many academics around the globe debated the impact of the digital gap in their respective communities with regard to local challenges (Preece & Rogers & sharp, 2015). One may wonder though, whether the gap is now closed in South Africa with the decreasing connectivity costs and availability of low-cost handheld computing devices (Oz & Jones).

Fitt's model was developed four decades ago to explore and understand the ability of new users to adapt to existing machinery. The theory that underpins this model is still valid in our lives today, especially when it comes to understanding the time it takes people to master computing systems, the rate of error, and the speed of processing once a task is repeated.

1.3 Prior research

Many computing usability studies have been conducted in different parts of the world. In most cases, usability testing focuses on average users and power-users, and less attention is paid to first-time users, especially those without previous access to any form of computing in their lives (Freeman & Freeman, 2005). These are often people living with high levels of poverty. It has therefore become essential that a unique study focuses on such people who can be considered as new entrants in the field of computer usage. This project is unique in that similar studies have not been done in this area that incorporate Fitt's model as well as the living conditions, levels of education, and the overall socio-economic conditions of users (Olivier, 2006).

1.4 Central research question

What are the usability challenges faced by new or first-time handheld device users given the current digital gap, and in view of the social, economic and education backgrounds of such users?

2 Method used (one example only, due to space constraints)

To demonstrate what a handheld computing device is and what constitutes new computing trends?

2.1 Research method

A qualitative research method was used in this section, motivated by the fact that the researcher needed to collect pertinent data on current trends of handheld computing business devices.

2.2 Data collection techniques

For this section, the researcher interviewed developers of computing systems, programmers, designers, and retailers.

2.3 Research population

The population was essentially ICT professionals who were involved in solving digital problems on a daily basis. The sample population consisted of individuals who sell computing devices and those who design, manufacture, and test handheld computing devices.

2.4 Research sampling

Probability sampling was used as a method of narrowing the sample down to the expected target group. This provided an equal opportunity to participants of the target population to participate in the project.

2.5 Research design

A quasi-experimental design approach was used.

2.6 Research location/area

In line with the main research objective, which was to conduct a nationwide research study on the usability of handheld computing devices, nationwide interviews were conducted with all participants.

2.7 Data collection techniques

Interviews were conducted and recorded at the participants' place of preference during the course of their working day.

3 Research hypothesis

Computing systems including handheld tablet systems manufactured in developed countries do not pay attention to usability challenges faced by first-time users in low-income communities in developing countries.

4 Expected contribution

This research will assess the current levels of usability of handheld computing devices for first-time users who are based in remote communities in South Africa. People in such communities who are less exposed to new technologies are the primary target of this research. The following will be taken into account – the ergonomic system, the relationship between human beings and computing, adaptation, error rates, task completion times, user satisfaction, and the cost of handheld computing devices in South Africa. Most importantly, the design of the user manuals of low-cost handheld devices will be assessed as well as their levels of user-friendliness, design, usability features, and cost.

5 Proposed solution

Closing the digital gap between those who have access to ICT devices and those who do not is a challenge for many governments around the world, regardless of continent, the state of the local economy, natural resources, or political policy or demographic make-up. The solution may not simply be the availability of, and access to, computing systems, but of effective and efficient usage of the handheld device. The goal is that there will be more new users that would later reduce or close the digital gap

It is anticipated that this study will contribute towards an understanding of the reason for the digital gap in South Africa, a situation that exists in spite of the availability of low-cost handheld devices. The researcher will apply Fitt's law to assess the success rates, task completion times, error numbers, and degree of frustration that first-time users experience. This will inform proposals for an innovative low-cost handheld computing framework specifically intended for the tablet that will contribute to the adaptation of handheld computing devices to first-encounter users in the future.

6 Statement to date

The following chapters and sections have either been completed or are in the process of completion:

- Chapter 1: completed
- Chapter 3: completed
- Chapter 2: completed
- Data collection: in progress.
- Faculty research ethics clearance: completed
- Questionnaires: completed.
- Data analysis: in view

7 Questions for discussion at the DC:

1. What is the importance of Human Computer Interaction (HCI) when it comes to computing?
2. What impact does HCI have on handheld computing systems usability?
3. How can computing systems designers balance systems learnability, efficiency, memorability, user satisfaction, and sound computing design principles?
4. What is Fitt's law?
5. What implications, if any, does Fitt's law have on computing systems adaptation?
6. What is STEA analysis?
7. What impact does it have on handheld computing?
8. What is the link between Fitt's law and STEA?

9. How should Fitt's law and STEA be calculated and applied to handheld devices?

References

[1] Freeman, E., and Freeman, E. Head first, a brain-friendly guide. O'Reilly. 2005.
[2] Lazar, J. Universal usability, Designing interfaces for diverse users, John Wiley & sons, Ltd, 2007.
[3] Lazar, J., Feng Heidi J,. Hochheiser, H. Research methods, in human computer interaction, John Wiley and sons, 2010.
[4] Preece. Rogers. Sharp. Beyond human computer interaction, John wiley and sons, 2002.
[5] Preece. Rogers. Sharp. Interaction design, beyond human computer interaction 4th edition, John Willey and sons, 2015.
[6] Olivier, M. Information Technology research, a practical guide for computer science and informatics, 2nd edition. 2006.
[7] Oz, E. and Jones, A., Management information systems, cengage, 2006.
[8] Reis, E., Usable usability, simple stepts for making stuff better, John wiley and sons, 2012
[9] Tegarden, D., Dennis, A and Wixon, B., systems analysis and design with UML 4th edition, John wiley and sons, 2013.

A Model for Parental Control Systems on Mobile Devices

Ibrahim R. Mbaya
University of South Africa
Pretoria, South Africa
36446106@mylife.unisa.ac.za

Abstract. With the prolific growth in usage of mobile devices by children and unique challenges in developing parental control systems for mobile devices, parents are looking for effective ways to control access to inappropriate contents by their children. Hence, this study investigates the existing models for parental control systems on mobile devices and assess their effectiveness. A model-building method will then be used to propose an effective model for parental control systems on mobile devices. This study will contribute by developing an effective model for parental control systems on mobile devices that integrates essential features derived from the established requirements of a parental control system to make parental control systems more convenient to users, i.e. both parents and children. A prototype will be developed to demonstrate a method of developing an effective model for parental control system on mobile devices.

1 Research problem

The existing parental control systems are weak to meet the expectations of most parents such as failures to filter harmful contents that was observed on Cyber Patrol, Cyber Sitter, Cyber Snoop, Net Nanny, and Surf Watch systems (Clayton, 2005; Keller & Verhulst, 2000). Parents want more transparency in their children's use of mobile devices (Yardi & Bruckman, 2011).

The main concern of parents is to strike a difficult balance between allowing their children to access information and communicate with their friends, while keeping them safe online (BAE Systems, 2012). Moreover, parents want more convenience on monitoring their children's internet activities so that they can monitor their children's internet activities while they are anywhere (Noor, Noor, Syed, & Zakaria, 2012).

The issue of effectiveness of parental control systems is an important component from a technical and fit for purpose point of view (Mielech, 2012). Researches on parental control systems show that existing parental control systems are not effective and have less user adoption rates (Hart Research Associates, 2011; Jigsaw Research, 2012; La Polla, Martinelli, & Sgandurra, 2013; Mielech, 2012).

2 Proposed research methods

The proposed approach for this study is referred as sequential exploratory strategy (Creswell, 2009), which is a mixed methods approach that involves a first phase of qualitative data collection and analysis through the literature survey, followed by a second phase of quantitative data collection and analysis through model-building study, in which effectiveness of existing parental control systems will be measured and an effective model for parental control systems on mobile devices will be developed.

Methods to deploy parental control systems on mobile devices will be extracted from the existing literature. A model-building study will be used to assess the effectiveness of existing methods and propose an effective method to deploy mobile services for parental control systems on mobile devices. Lastly a prototype will be used to develop a parental control system that uses the improved effective model for parental control systems on mobile devices. The prototype will be used with parents/children to validate is effectiveness and thus evaluate the research work.

3 Expected contributions

The main contribution from this study will be an effective model for parental control systems on mobile devices that integrates essential features of a parental control system. This will make parental control systems on mobile devices more convenient to users, i.e. both parents and children.

4 Statement of work to date and open issues

The work that have been done already includes the preliminary literature survey that justify the research problem and establishing the research methods to carry out the research.

4.1 Findings from preliminary literature survey

The preliminary literature survey indicated that existing parental controls features monitoring, filtering, alerting, and reporting of access to information and resources on mobile devices. Additional aspects of existing parental control systems are location-aware and context-aware. These features are summarised below.

Parental control model	Monitoring	Filtering	Alerting	Reporting	Location-aware	Context-aware
Debbabi et al. (2007)	✓	✓	✓			
Behrooz & Devlic (2012)	✓	✓				✓
Kuppusamy et al. (2013)	✓	✓	✓	✓	✓	
Noor et al. (2012)	✓	✓				
Costa et al. (2010)	✓	✓			✓	
Decker (2011)		✓			✓	
Thierer (2007)	✓	✓	✓	✓		
Fellenstein et al. (2007)	✓	✓		✓		
Gatz et al. (2002)	✓	✓				
Parker (2014)	✓	✓			✓	✓
Matthews et al. (2013)		✓	✓		✓	✓
Ackley (2007)	✓	✓		✓		

Table 1. Preliminary analysis of features of existing parental control models

From the analysis done on the literature it was found that the existing models of parental control systems on mobile devices does not integrate all these features.

4.2 Open questions/issues

From the preliminary literature survey, several questions/issues remained to be answered. The list of open issues includes:-

- What are the requirements of effective model for parental control systems on mobile device?
- How do we measure the effectiveness of a parental control system on mobile devices?

These open issues are expected to be resolved through the research methods proposed above.

References

[1] Ackley, J. (2007). Cell phone parental control. USA. http://www.google.com/patents/US7302272

[2] BAE Systems. (2012). Detica StreamShield Parental Controls. Sundby: Global Communications Solutions. www.baesystemsdetica.com/gcs

[3] Behrooz, A., & Devlic, A. (2012). A Context-Aware Privacy Policy Language for Controlling Access to Context Information. Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, 25–39.
[4] Clayton, R. (2005). Failures in a Hybrid Content Blocking System. In Privacy Enhancing Technologies (pp. 78–92). Cambridge: Springer Berlin Heidelberg. Retrieved from http://dx.doi.org/10.1007/11767831_6
[5] Costa, G., Torre, F. la, Martinelli, F., & Mori, P. (2010). Parental Control for Mobile Devices. European Research Consortium for Informatics And Mathematics, 82.
[6] Creswell, J. W. (2009). RESEARCH DESIGN: Qualitative, Quantitative, and Mixed method Approaches (Third Edit.). Lincoln: SAGE Publications.
[7] Debbabi, M., Saleh, M., Talhi, C., & Zhioua, S. (2007). Embedded Java Security: Security for Mobile Devices. Zhurnal Eksperimental'noi i Teoreticheskoi Fiziki. London: Springer Science & Business Media.
[8] Decker, M. (2011). Modelling of Location-Aware Access Control Rules (pp. 912–929). IGI Global. doi:10.4018/978-1-60960-042-6.ch057
[9] Fellenstein, C. W., Gusler, carl P., Hamilton, R. A., & Schatz, H. (2007). Apparatus and method for monitoring and analyzing instant messaging account transcripts. USA.
[10] Gatz, S., Guggenheim, B., Lin, T., Robison, J., & Toth, A. (2002). Parental control system for use in connection with account-based internet access server. USA. http://www.google.com/patents/US20020049806
[11] Hart Research Associates. (2011). Who Needs Parental Controls? A Survey of Awareness, Attitudes, And use of Online Parental Controls. New directions for youth development (Vol. 2013, pp. 5–8). doi:10.1002/yd.20075
[12] Jigsaw Research. (2012). Parents' views on parental controls: Findings of qualitative research. http://stakeholders.ofcom.org.uk/binaries/research/media-literacy/oct2012/Annex_1.pdf
[13] Kuppusamy, K. S., Mary Francis, L., & Aghila, G. (2013). Report: A Model for Remote Parental Control System Using Smartphones. International Journal on Cybernetics & Informatics, 2(3), 25–36. doi:10.5121/ijci.2013.2303
[14] La Polla, M., Martinelli, F., & Sgandurra, D. (2013). A Survey on Security for Mobile Devices. IEEE COMMUNICATIONS SURVEY & TUTORIALS, 15(1), 446–471.
[15] Matthews, J. H., King, J. S., Schrader, J. A., Chen, T. T.-Y., & Sarin, R. K. (2013). Automatically quieting mobile devices. USA. http://www.google.com/patents/US20130225152

[16] Mielech, N. (2012). Benchmarking of parental control tools for the online protection of children SIP-Bench II. Safer Internet Programme. http://www.sipbench.eu/transfer/SIP_BENCHII_4th_cycle_report_FINAL.pdf

[17] Noor, R., Noor, S., Syed, S., & Zakaria, K. H. (2012). Parental Mobile Control System for Children' s Internet Use. International Conference on Information Society,511–513. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6285036&isnumber=6284715

[18] Parker, S. B. (2014). Child/User mobile device manager. USA. http://www.google.com/patents/US20140171059

[19] Thierer, A. (2007). Parental Controls & Online Child Protection: A Survey of Tools and Methods. Washington, D.C.: The Progress & Freedom Foundation. www.ppf.org/parentalcontrols

[20] Yardi, S., & Bruckman, A. (2011). Social and Technical Challenges in Parenting Teens' Social Media Use. In CHI2011 (pp. 3237–3246). Vancouver: ACM. doi:978-1-4503-0267-8/11/05

Workshop:
IFIP WG 13.5 Workshop on Resilience, Reliability, Safety and Human Error in System Development

Alignment of Technology to Work: Design & Evaluation Representation

Dorrit Billman[1], John Archdeacon[2], Rohit Deshmukh[1], Michael Feary[3], Jon Holbrook[1], Michael Stewart[1]
[1]San Jose State University, [2] Dell, [3]NASA Ames Research Center
San Jose State University @NASA Ames, Moffett Field, CA 94035-1000
dorrit.billman@nasa.gov

Abstract. Robust and effective technology for high-stakes work depends on an appropriate understanding of the work needs. We summarize a method for representing work, technology, and their alignment, which aims to guide technology design and evaluation and to ensure that technology is fit-for-purpose. We illustrate with aerospace examples.

1 Introduction

Effective integration of humans and technology in safety critical systems requires an understanding of what functions the system is intended to support. Work needs should drive a) design, for example by driving requirement specification, and b) evaluation, for example through validation. When analysis of work needs guides design and evaluation, this should ensure effective, robust technology for carrying out the functions needed across a work domain.

In the domain of home entertainment, one intended function is recording for later viewing an upcoming event, such as the Singapore Grand Prix. Some technology designs may require the user to seek elsewhere information about time and local broadcast source, then specify time and channel on VCR, command the recording, and "manage the automation" by checking space limitation or other possible conflicts. A device better fit-for-purpose would provide all the information and operations for each function, accessible together. Issues of fitness-for-purpose can arise in any activity domain but are particularly important for high stakes and safety critical work.

The functions that the system should support might only emerge over the course of development and evaluation. It would be efficient, however, if the analysis preceded design and development, so missing or misappropriated functions are not encountered only in evaluation. Identifying and supporting work needs early in the design process prevents expensive redesigns and expensive training to compensate for a design not fit-for-purpose. A method

for identifying work needs early and a representation that can be shared and used throughout the design-to-delivery phases is very desirable.

This paper describes a method and representation of work, of technology, and their alignment, which we are currently applying and evaluating. The representation and method aim:

- to guide development of requirement specifications (or analogous documents) particularly in high-stakes work domains;
- to guide design; and
- to provide for ongoing evaluation, by assessing alignment of technology designs with the work.

Further, providing a representation of the work to be done provides input to the function allocation policy determining what agent does what components of work. When requirement specification is guided by work needs analysis, the requirements can provide a stable, shared representation throughout the specify-design-evaluate process, e.g. as in [1], that keeps the process anchored to actual needs of the work. Multiple researchers have noted the value of work representations and we have been influenced by many of these [2]–[7].

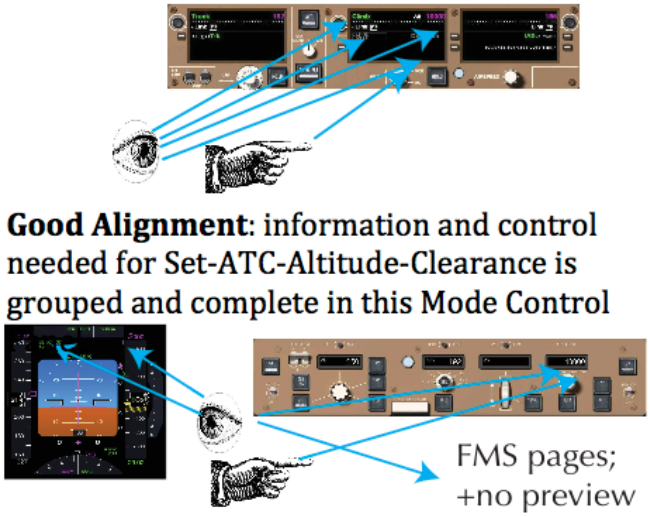


Figure 1. Illustrative example of good and poor alignment of cockpit elements use for tactical change in course, typically compliance with ATC clearences.

We claim that fitness-for-purpose depends on the alignment of the structure of the technology with the structure of the work. By structure we mean the components and their organization, of the work and of the technology. By alignment we mean that the components of work map relatively directly onto the components of technology. In turn this facilitates transparency of the technology, so that workers can focus on the work they wish to accomplish rather than the needs of the technology they are using to get the work done. For example, if the work domain is piloting autoflight-capable jets, a cockpit

design that is well aligned with the work will be organized around flight rather than the internal workings of the autoflight system. Of course, management of the technology (the autoflight system) becomes part of the work (flying an airplane), but well-aligned technology adds minimal and subsidiary demands. Figure 1 illustrates how technologies might differ in alignment for a single work function -- complying with an Air Traffic Control (ATC) clearance to change altitude. This requires as input variables the present altitude, the current target altitude, the altitude specified by ATC, and the current autoflight mode; it changes as output the current target altitude and possibly the mode. Alternative Mode Control Panels, as sketched, may be better or more poorly aligned with this function.

A key insight in developing our representation was recognizing that the input and output variables relevant to work functions can act as a "common language" to describe both the components of work and the components of the target technology. This common language mediates the assessment of alignment of technology to the work. In overview:

- The work and the technology are each represented as a matrix.
- The work matrix has the intended work functions listed as one dimension and the variables used or affected by the work on the other dimension.
- The technology matrix has the technology components listed as one dimension and the variables that the technology provides through displays and controls as the second dimension.
- Alignment can be assessed by comparing the components and structure of the technology matrix to the work matrix.

2 Work Functions, Variables, and Matrix

We represent work in terms of a set of intended *work functions* and a set of *variables,* within a bounded *work domain*. A *work domain* is a body of related activities intended to accomplish a set of stable goals, such as making products or controlling processes.

An intended *work function* is an activity needed to accomplish the mission or achieve the goals of a work domain. To be the most useful guides for technology design and evaluation, work "functions" are more abstract than what are typically referred to as "tasks." New technology can transform work, and work functions should be written at a level sufficiently abstract to allow variation in how the same function is to be accomplished across alternative technology designs. Granularity of the work functions is an important analysis choice, influenced by the analysis purpose and the scope of the technology;

more specific breakouts of functions might be undertaken within a larger analysis as well.

Work functions are represented in terms of the variables needed as input to and the variables affected as output from the function. Input variables represent the information and resources needed to do the work. For cognitive work, information variables are typically the key concern in designing and evaluating technology, that is, much of the design challenge concerns how to make available all the information needed to make decisions without producing information overload. Output variables are changed as the result of carrying out a work function. Output variables broadly map onto the aspects that can be affected by any control action. Output variables can change the state of the interface, the state of the system behind the interface, an information variable such as a decision, or the information made available to different agents, as through calculation or communication. Identification of the *input* variables needed across a work domain has also been stressed by developers of ecological interface design, as the *information requirements* [8]. We also recommend a parallel census for output variables. In the aviation case we have looked at in most detail, the number of input variables seems to be substantially higher than the number of output variables.

Given a set of work functions and a set of the input and output variables, each work function is coded for what information variables it needs for input and what output variables express the changes accomplished by the function. Consider the function of executing a highly planned attitude manoeuver of the International Space Station (ISS) to accommodate a Soyuz docking. Input variables would include the systems to be monitored to ensure that the planned and actual situations align: S-band communication availability, status of the momentum management system on the ISS, status of controlling software on the ISS, trajectory of the approaching Soyuz, status of the software controlling Soyuz, status of crew, and status of interrelated systems on the ISS such as configuration of the solar arrays. Output variables would include issuing commands to the momentum management system and providing information to ground controllers about progress. Some variables may be complex and derived as the output of other work functions, such as a summary representation that a particular system is ‘normal’. Our concept of work functions with input and output values is quite similar to Pritchett et al’s [9] representation of work functions.

Given a census of work functions and of the input and output variables for a work domain, the relation between these can be expressed in a matrix. A work function can be represented as a vector of variables, with 1s indicating a variable is relevant to the function, and 0s otherwise. Figure 2-A illustrates

the individual work function vectors integrated into a matrix. The top line in the matrix represents work function 1 as the row vector (0,0,0,1,0,0,0,1,1,0,0,0,1). Note this vector does not specify *what* the value of the variable is, but simply that the variable is important for carrying out that function. Conversely, a variable can be represented as a column vector indicating what work functions use that particular variable, the vector (0,0,0,0,0,0,0,0,1,1) for input variable 1.

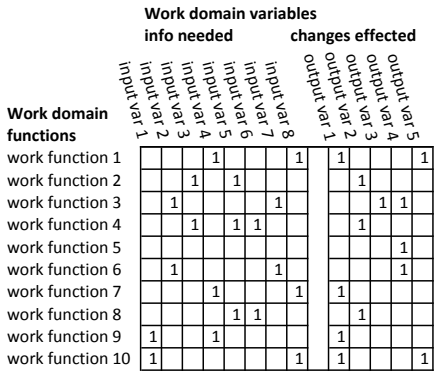


Figure 2A: Work Matrix

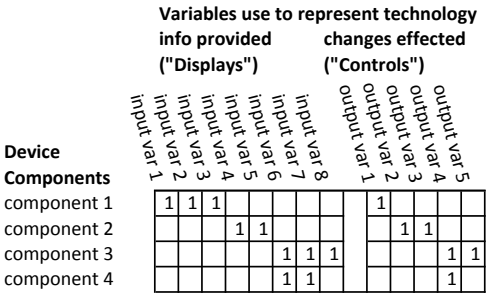


Figure2B: Technology Matrix

3 Technology Components

Just as a work domain can be decomposed into component work functions, a complex technology can be decomposed (or built up) from its components. A component is a group of interaction elements, i.e., a set of displays and controls, related by spatial or temporal proximity. Both spatial (e.g., located close together or within the same physical unit) and temporal (e.g., information appearing on a screen at the same time) proximity of elements are important in interaction [5] and in setting component boundaries. Temporal proximity and distance is relevant when the interaction options change with time, enforcing sequential steps to access; for example, requiring navigation through multiple screens to access one from the other or if mode changes restrict what resources are available at one time. Spatial proximity is more studied in automation [10], and it would be unusual to group together as one component, elements that are widely separate spatially. We believe elements of a complex automation system are often marked jointly by spatial proximity, simultaneous availability, and development or manufacturing history. However, as with identifying the domain work functions and variables, picking the size, or level, of the components requires judgment and will depend on the purpose of the analysis. For design, identifying higher-level

parts initially, and adding a detailed analysis of selected components may be helpful. For evaluation, definition of lower level components may be helpful. For simplicity we assume that components at the within one analysis are disjoint.

The same variable set is used for representing technology components as for representing work functions, providing a common language. Just as a work function can be scored for which variables are relevant, a technology component can be scored for the displayed information (providing input to the user) and controllable variables (enabling action by the user) it provides. Note that the technology *displays* provide work function input-variables and the technology *controls* provide work function output-variables. As with the vectors for individual work functions, an individual technology component can be scored as a vector of 1's and 0's showing the variables it supports. Again, this vector does not specify *what* the value of the variable is that the component provides, but simply that the variable is provided by the component. Conversely, a variable can be represented as a vector indicating what components provide support for it. Typically, components also include displays and controls that do not have to do with the work domain but only with operation of the technology. These can be thought of as "overhead"; typically the variables used to represent the components will be the union of the domain and the overhead variables.

Just as the work domain can be represented as a matrix, so can the technology. The matrix representation of work and technology, with shared dimension of variable, supports analysis of alignment.

4 Alignment of Technology with the Work Domain

There are three aspects of alignment: coverage, overhead, and organization. **Coverage** refers to the proportion of work domain variables that are provided somewhere in the technology. Technology with low coverage cannot go very far supporting the breadth of work functions. While it may not always be desirable or possible for technology to provide maximal coverage (some variables may be better left to a person), it is valuable in any case to have an explicit accounting of what variables the technology is and is not supporting. Technology with higher coverage is better aligned and typically better able to support its intended purpose, ceteris paribus.

Overhead refers to the variables that are not concerned with getting domain work done, but exclusively concern operation of the technology that is the focus of the design or evaluation. The definition of overhead critically depends on the scope of the work domain. If the work domain is work of the cockpit

crew piloting an autoflight-capable jetliner, the work functions describe what is needed to fly the jetliner on the intended trajectory with appropriate communication; the overhead measures the work needed to manage the devices in the cockpit such as switching among modes or accessing information in the flight plan. Technology with high overhead may be "rude" or akward to use. Technology with lower overhead is better aligned, and typically better able to support its intended purpose.

Organization refers to the way that the work functions and the device components group and are grouped by the variables. In a simple domain, there might be a one-to-one mapping between work functions and technology components. Every work function would have its own technology component, providing support for all the relevant variables, whether a separate web page or a separate, specialized tool for a physical assembly job.

In a slightly more complicated situation, several work functions might be accomplished with the same, single component, producing a many-to-one mapping. For complex work the "best" design may provide a many-to-many mapping: there may simply be too many combinations of variables needed by different work functions to make a one-to-one mapping feasible. However, mapping clusters of related variables onto a component can still produce a well-aligned organization: the variables grouped together in one component are organized to provide most of the variables needed by most of the functions in that cluster. If there is little similarity structure across work functions—each picking a very different set of variables from the others, the is a very limited possibility for aligned organization. The technology design could pick a few high-stakes or frequent work functions, design to make support for these few coherent (for however much complexity can be tolerated) and leave the remaining functions to "piece together" their needed variables, which are scattered across components.

Some aspects of organization can be "read off" from a matrix through sorting and inspection. Disjoint sub-matrices may reveal clearly separable groups of functions, which share variables within but not between function groups. However, recovering the implicit matrix organization is difficult to do by inspection and we have been developing and applying (bi) clustering methods, focusing on discovery of structure within the work function matrix.

A preliminary analysis of alignment was developed for aspects of the aviation domain, particularly for an airline pilot's tactical changes in trajectory in response to Air Traffic Control clearances.

5 Exploratory Evaluation: Effect of Alignment on Ease of Learning

Rationale. Technology that is better aligned with a work domain (or part thereof) should be easier to learn for domain experts. That is, the more that working *through* the technology is based on work activities (already familiar to domain experts), the less that operation requires learning new, technology-specific skills and knowledge. One aspect of alignment is whether, work function by work function, the technology provides displays for the information variables and controls for the output variables for each work function; and whether the interface groups together those elements that are needed together.

Method. We identified and evolved two designs for Mode Control Panels (MCPs). MCP's are intended for use by an airline pilot to enter tactical changes in the airplane's trajectory, typically in response to Air Traffic Control clearances. These two MCP designs differed in how directly they were aligned with the needs of work functions for complying with ATC clearances. One important difference between the two MCP designs was the organization of displays and controls in the interface. In the cockpit design with the *Alpha* MCP the displays and controls for work functions such as compliance moving to a newly specified altitude were distributed across several cockpit devices (the MCP, the Primary Flight Display, the Flight Management System); in the cockpit design with the *Bravo* MCP the displays and controls for such ATC clearances were integrated within the MCP, hence spatially close and available without paging or navigating. Thus, we anticipated that learning the particular actions needed to implement ATC clearances would be faster in a cockpit using the Bravo rather than Alpha MCP, for 1) pilots needed to implement ATC clearances would be faster in a cockpit using the Bravo rather than Alpha MCP, for 1) pilots familiar with ATC clearances and flying in the National Air Space but 2) unfamiliar with either MCP design.

We developed 8 scenarios, flown in a mid-fidelity simulator (Figure 3). Six involved vertical and speed changes (where we anticipated most difference) and 2 involving lateral changes. We used a trials-to-criterion learning procedure with a criterion of 2 successive correct executions of a particular scenario. The scenarios were repeated, in sequence, with individual scenarios dropping out as the participant reached criterion on that scenario. Users were corrected when they made an error and the scenario stopped. Participants were 6 regional pilots who had experience with autopilot aircraft, but were screened to be unfamiliar with the Alpha interface (similar to one in commercial use).

Follow-up questioning found that User #3 (in Figure 4) unfortunately did in fact have experience relevant to the Alpha interface.

Figure 3. Current simulator cab with lab-designed Prototype Design. Participant sits at left (Captain's seat) and experimenter at right (First Officer seat). Experimenter uses tablet computer to control progress through the experiment.

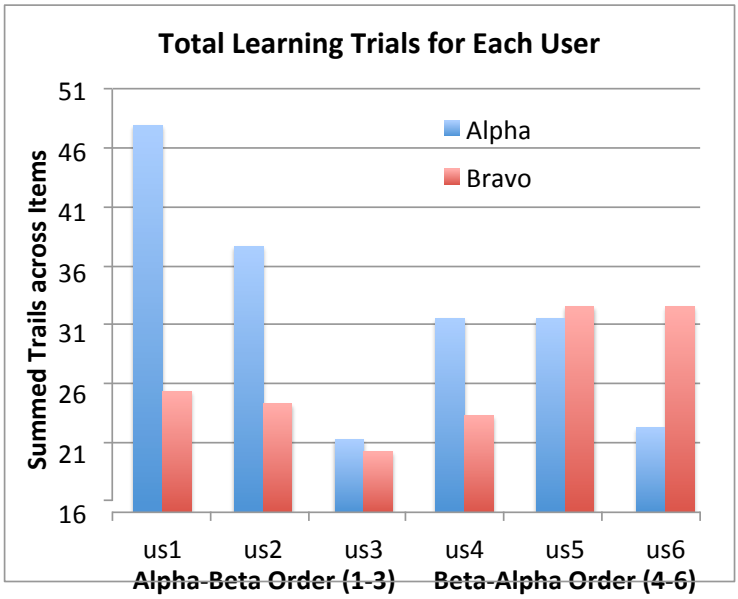


Figure 4. Trials-to-criterion with Alpha and Bravo MCP designs. Users are grouped by which MCP was used first. Users 1-3 used Traditional first; users 4-6 used Prototype first. The figure suggests better learning on the second interface and better learning for Alpha, substantiated in analyses.

Results and Analysis. Figure 4 shows trials to criterion for cockpits with each MCP design, grouped by which system the participant learned first. We modelled the factors influencing trials-to-criterion with linear mixed model regression, which offers a sensitive analysis for within-subject and small N designs. We used R (R Core Team, 2012) and *lme4* [11], with the

untransformed trials-to-criterion as the variable predicted. Fixed effects in the model were MCP Condition (Traditional vs Prototype), Order (First vs Second experience), and Item. Subject (intercept only) was a random effect. Using the *Bravo* MCP design reduced the learning trials by 0.71 trials (StDev = 0.2716, associated t= -2.608). To test the significance of the effect of Condition, we used the maximum likelihood approach comparing this model to the null model from which Condition had been dropped, and found that Condition is a significant factor ($\chi_2(1)=6.557$, p=0.01045). An analogus analysis showed that Order was also significant ($\chi2(1)= 8.9537$, p=0.002769); users were faster learning the second system they were introduced to, with this reducing the trials to criterion by 0.833 trials (StDev=.2716, t=-3.068).

This preliminary, low n study provided suggestive evidence for ease of learning of the Bravo over Alpha design, using a trials-to-criterion learning measure. Limitations include small n, few scenarios, use of one learning method, and imprecision in scope of the work functions assessed. Nevertheless, this study suggests that the alignment of the technology's content and organization to the work domain's content and organization can be assessed and that alignment may be important for learnability.

6 Conclusions & Future Research Needs

To ensure resilience of safety-critical systems, the integrated performance of the human-plus-engineered components must reliably support the functions needed in the work domain. This requires understanding the work, the technology, and how the technology is related to the work. Each of these is complex, and there are complementary perspectives for addressing complexity: depth first or breadth first. A 'depth first' perspective focuses on a sample of key elements in the work and in the technology. Each of the sampled scenarios may be evaluated in detail, but it may be infeasible to assess the whole work domain in such depth. A 'breadth first' perspective focuses on covering all aspects of a well-defined work domain and candidate technology, though in less detail. In our 'breadth first' approach, we use a relatively simple representation of both work and technology. A "common language," for the informational variables needed to make decisions and the action variables needed to effect changes, can describe and compare the work functions and the technology support at a relatively abstract level. An important research goal will be exploring these complementary perspectives and, particularly, how they can be usefully coordinated.

Acknowledgements

This research was funded in part by the NASA Human Research Program: Space Human Factors Engineering (466199.02.01) PI, Michael Feary. Thanks to members of the AID lab group. Thanks to Emilie Roth for comments on an earlier draft.

References

[1] R. W. Pew and A. S. Mavor, Eds., *Human-System Integration in the System Development Process*. Washington, D.C.: National Research Council, The National Academies Press, 2007.

[2] K. A. Butler, J. Zhang, C. Esposito, A. Bahrami, R. Hebron, and D. Kieras, "Work-centered design: a case study of a mixed-initiative scheduler," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, 2007, pp. 747–756.

[3] K. Holtzblatt, J. B. Wendell, and S. Wood, *Rapid contextual design : a how-to guide to key techniques for user-centered design*. San Francisco: Elsevier/Morgan Kaufmann, 2005.

[4] H. Johnson and P. Johnson, "Task knowledge structures: Psychological basis and integration into system design," *Acta Psychol. (Amst.)*, vol. 78, pp. 3–26, 1991.

[5] D. Kieras, "Task analysis and the design of functionality," in *The Computer Science and Engineering Handbook*, vol. 23, A. Tucker, Ed. Boca Raton: CRC Inc., 1997, pp. 1401–1423.

[6] N. Naikar, R. Hopcroft, and A. Moylan, "Work Domain Analysis: Theoretical Concepts and Methodology," Australian Defence Science and Technology Organisation, Melbourne, Australia, 2005.

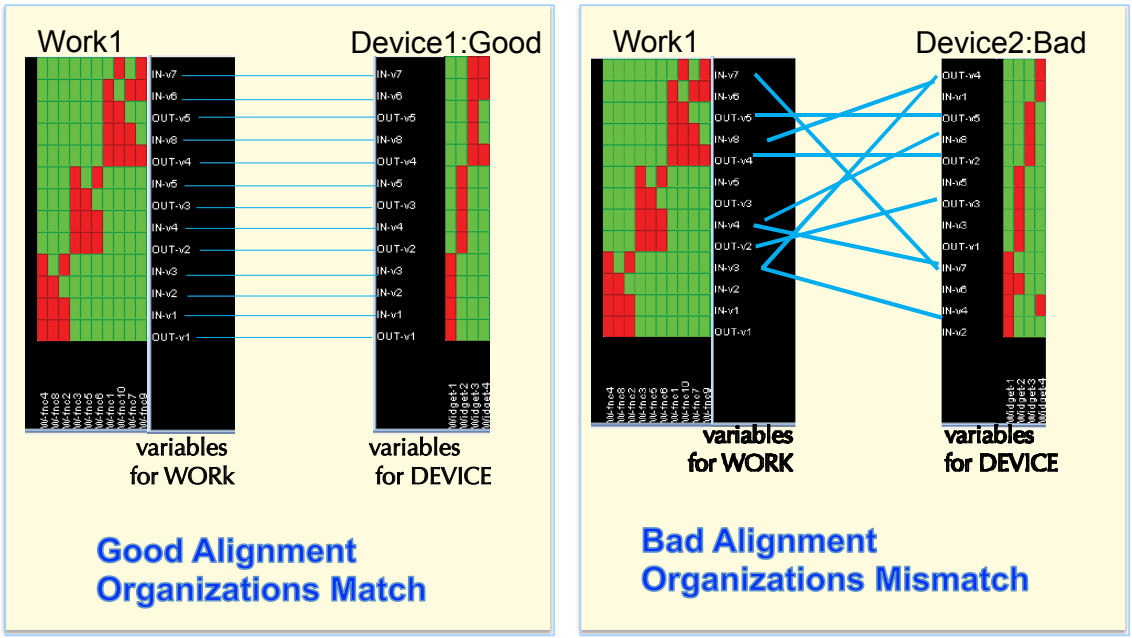
[7] E. M. Roth and R. G. Eggleston, "Forging New Evaluation Paradigms: Beyond Statistical Generalization," in *Macrocognition Metrics and Scenarios Design and Evaluation for Real-World Teams*, E. S. Patterson and J. E. Miller, Eds. Surrey, England: Ashgate Publishing, 2013.

[8] C. M. Burns and J. R. Hajdukiewicz, *Ecological interface design*. Boca Raton, FL: CRC Press, 2004.

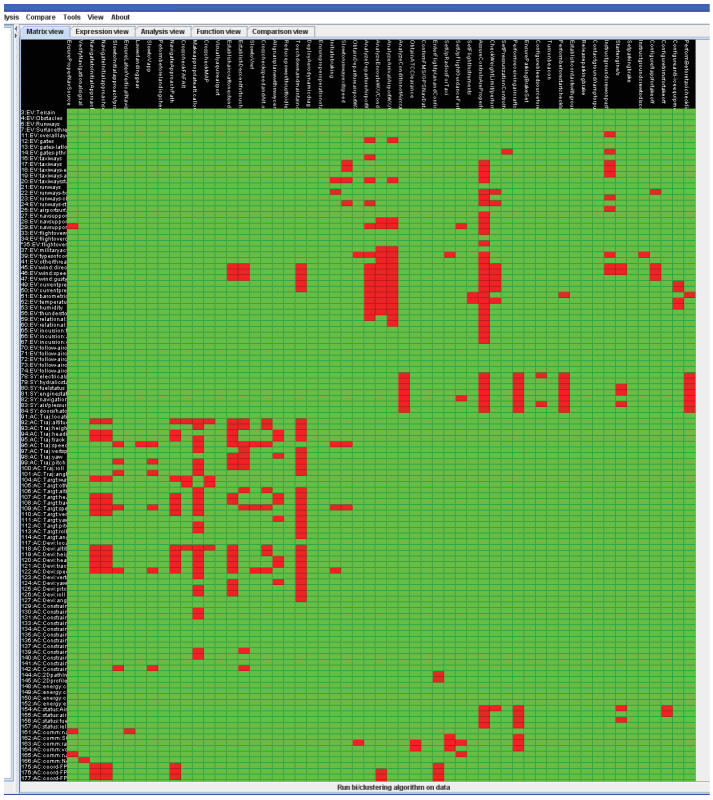
[9] A. R. Pritchett, S. Y. Kim, and K. M. Feigh, "Measuring Human-Automation Function Allocation," *J. Cogn. Eng. Decis. Mak.*, vol. 8, no. 1, pp. 52–77, Mar. 2014.

[10] C. D. Wickens and A. D. Andre, "Proximity Compatibility and Information Display: Effects of Color, Space, and Objectness on Information Integration," *Hum. Factors J. Hum. Factors Ergon. Soc.*, vol. 32, no. 1, pp. 61–77, Feb. 1990.

[11] J. Pinheiro, D. Bates, S. DebRoy, and D. Sarkar, *Fit and compare Gaussian linear and nonlinear mixed-effects models*. R Package Version nlme 3.1-118, 2014.



APPENDIX *1 Alignment Illustration*. The left panel shows a technology perfectly aligned with the work functions. The technology groups displays and controls for these variables to match the grouping based on work functions. The right panel shows bad alignment. Some variables in the work domain are not covered, the device includes overhead variables, and the way the displays & controls group variables is unrelated to the way the work functions group variables.



APPENDIX 2. Part of the 119 x 210 function X variable matrix for cockpit aviation. Inspection suggests loose cluster structure, where sets of work functions draw on sets of common variables. The visual patterns suggest the relevance of statistical clustering methods to identify domain structure .

Reflecting on Users' Strategies for Resilient Interactions

Jonathan Day, George Buchanan, Stephann Makri
Centre for HCI Design, City University London
Northampton Square, London, EC1V 0HB, UK
Jonathan.Day.2@city.ac.uk

Abstract. One crucial contributor to the resilience and reliability of interactions with technical and sociotechnical systems is the resilience of users themselves. While the study of human factors has traditionally focused on the negative aspects or frailties of human performance, attention is increasingly turning to also consider the proactive and positive contributions human performance can make across a range of tasks and settings. In this position paper, we introduce the notion of *Resilience Strategies*, summarise some of our current work in this area and discuss examples of resilience strategies we have encountered during the course of this work. We also discuss how work into resilience strategies is situated in terms of broader work into the high-level resilience of sociotechnical systems, and interactions with technical systems.

1 Resilience as Individuals' Behaviour

Resilience Engineering has, to date, largely represented the pursuit of heightened resilience at a system or organisation wide level. At the same time however, a recent resilience-focused literature review [1] notes how resilience can be considered as a concept that scales down from high-level organisations, to groups and teams, and further still to an individual level. As an illustration of this, resilience could for example be considered in (i) the way in which emergency services respond to a large-scale incident (high level 'systems' resilience) and/or (ii) in the way resources are allocated by a team of dispatchers (the resilience of a small team within the system), and/or (iii) in the way in which an ambulance driver may for instance use conscious foresight to select a longer but faster or less traffic-prone route when moving to an objective (resilience demonstrated by an individual through the implementation of a resilience strategy).

While we acknowledge the valuable insight that can be gained from work into resilience at higher levels of granularity, in order to situate and contextualise our current perspective we note here that our work principally addresses the later of these levels; the resilience of individual operators or users. We propose that resilience at this individual level is perhaps most approachable when conceptualised in terms of *resilience strategies*, the

tangible behaviours and tactics that individuals deploy to mitigate threats and maintain performance.

Strategies leading to improved performance through resilient actions and interventions are nothing new, and examples are observable across a range of contexts. One challenging aspect of their investigation however is that as with the wider study of resilience, such strategies are generally underrepresented in the literature owing to a tendency to instead direct investigative scrutiny towards failures, adverse incidents and threats or frailties [2]. There are however some cases, particularly involving work of a safety critical nature, where examples have been reported that capture individuals deploying strategies and workarounds to remain resilient against threats.

Randell and Johnson [3], for example, noted how they witnessed hospital nurses 'tricking' portable monitor devices by removing and reinserting batteries. This overcame a feature whereby every fiftieth charging cycle would trigger a battery condition error, necessitating battery replacement regardless of remaining capacity. In actual practice, it was impractical to record charging cycles and thus predict when this arose, but nurses could utilise this battery-removal strategy when the error presented at inopportune moments. Mumaw et al. [4] describe a number of strategies that operators deploy in nuclear power plant control rooms, for example reducing superfluous alarms by adjusting threshold parameters (increasing the salience of more important alarms), or the strategic repositioning of physical items to serve as visual cues for assisting with prospective memory and progress-tracking. Malakis and Kontogiannis [5] similarly discuss further such examples of strategy use observed in the work of air traffic controllers.

Furniss et al. [6] note that while targeted work addressing resilience and other closely related subjects is available, the tendency for this work to be presented across differing levels of granularity and abstraction, and generally within specific and specialised domains, means it can be difficult to assimilate such phenomena and transfer findings between domains. In response, Furniss et al. outline the *Resilience Markers Framework*, which seeks to explore how common themes may be derived from individuals' concrete strategies, and traced vertically to broader strategy-type patterns of behaviour, and ultimately high-level resilience markers or principles that are transferable across domains.

Furniss, Back and Blandford [7] developed this idea at an individual-strategy level by establishing a categorisation scheme for different types of resilience strategy, which they derived from a basic thematic analysis of some 49 episodes of self-reported resilience, collected expressly for the purpose. Furniss et al. arrived at a seven-item scheme that provides structure for further

analysis in the form of a vocabulary for describing different strategy types. However, the authors note that this scheme may potentially benefit from further refinement and validation, owing to ambiguities in terms of 'overlapping' in their categories and potential gaps in coverage reflecting their reliance on a somewhat limited dataset.

2 Refining Categories for Resilience Strategies

Part of the work we are undertaking seeks to extend and develop the aforementioned Furniss et al. categorisation scheme, in collaboration with the original authors. To date, this has involved the collection and analysis of an extended set of resilience strategies, combining multiple methods including observations from a controlled lab study, a diary study and self-reported episodes collected via a purpose-designed mobile application (an enriched 'digital diary' with prompts to elicit relevant contextual information, and the ability for users to upload photos, further described in [8]). We have further supplemented this data with strategies observed and noted from two additional studies by colleagues, which comprised situated observations of home healthcare, and an online survey into habits surrounding medication adherence.

As with the previous Furniss et al. investigation, and in a marked departure from much of the previous work into resilience strategies, we have opted not to limit the scope of our data collection to specific domains or tasks but to instead consider a broad spread of strategies from a variety of contexts. This reflects our objective of deriving a set of descriptors for types of resilience strategy which are independent of, and transferable across, domains and settings.

While this work is currently still in progress, this expanded dataset and subsequent analyses (particularly of 'edge-case' episodes of reported resilience) have already resulted in refinements to the scheme, and an increase in coverage and the total number of category descriptors. While the precise nature of much of this work extends beyond the current paper, we present here some examples of resilience strategy episodes we have collected as part of our data and discuss these in the context of the strategy categories they represent.

2.1 Examples of Individuals' Resilience Strategies

One such example of a resilience strategy was a diary study participant describing how they have purchased multiple chargers for their mobile phone. They either carry a spare with them, or leave it at their place of work. In so

doing, the participant is able to keep their phone charged while reducing the risk of a charger being forgotten. An indirect benefit of this redundancy would also be the continued availability of a charger in the eventuality of one becoming lost or developing a fault. We felt this strategy episode to be representative of the category descriptor *maintaining resource availability* and noted how strategies that primarily addressed this objective were a recurring theme in our dataset.

A second example of a resilience strategy, also recorded by a diary study participant, described how the participant had strategically customised the interface of a frequently used software package, reconfiguring toolbars for the purpose of reminding them about software functions that are otherwise hidden behind a menu. We characterised this as an example of the category descriptor *creating new cues*, as it represented an adaptation which utilises a visual cue to assist the user in avoiding a perceived potential memory slip error. We note that there are clear parallels between this onscreen cueing strategy, and the previously discussed Mumaw et al. [4] strategy describing how operators in a control room utilised physical artefacts as improvised cues.

A third example, which was reported by multiple participants across a range of contexts, described how smartphone users are increasingly utilising the camera functionality of their devices to quickly and conveniently capture a variety of information. Specific examples included capturing the contents of a whiteboard during a university class, generating a digital copy of a tube map, and transferring the contents of a digital calendar from a laptop screen to a tablet when the ability to synchronise was temporarily unavailable. We consider such examples to fall under the descriptor of *appropriating a resource,* since each case describes an improvised and unconventional use of the camera to generate externalised representations, which serve to offload working memory capacity.

Other such examples of types of strategy include *reinforcing an existing safety barrier* which can be observed when, for example, individuals set multiple 'back up' alarms to reinforce their effectiveness, *checking before or after an action* which describes checking variables prior to or following a task or subtask, and *separating task items* which involves disambiguating similar items or streams of information, for example the labelling or physical separation of two hard drives.

3 Implications for the Broader Study of Resilience

Returning to the broader investigation of the resilience of technical and sociotechnical systems from a HCI standpoint, while the 'micro-level'

accounts of resilience presented above may initially appear somewhat disconnected, we believe there is clear potential merit in the approach we are undertaking of adopting these concrete observable strategies as a base unit of analysis from which to derive more transferable principles at a higher level of abstraction.

These categories, originally proposed by Furniss et al. and currently the subject of refinement and validation, help us to articulate both the needs of users, and the resulting behavioural mechanisms by which they themselves proactively improve performance. We posit that by better understanding and nurturing these strategies, designers may in future be able to craft more resilient interactions, in turn improving performance and resilience in the wider encapsulating system.

We see two potential and complimentary avenues by which further insight of individuals' resilience strategies could be realised in the designs of future systems. Designers and practitioners could harness strategies for potential feature inclusions, using them to identify specific needs on the part of users. A HCI relevant example of this is the integration of automated checking in some modern email clients to prompt the user if an outgoing message is intended to be sent without an attachment, if the system detects the word 'attachment' in the body of the message. Alternatively, and perhaps more challengingly, designers and practitioners could endeavour to design-in capacity or flexibility to accommodate for, and facilitate, users developing and deploying their own strategies. One rudimentary example of this is the functionality in some file managers for arbitrary colour-coding or metadata fields for users to appropriate.

4 Conclusions

The pursuit of heightened resilience and reliability within complex systems represents a significant challenge. One route to achieving this might be through promoting the resilience of users at an individual or cognitive level. Variability in human performance has traditionally all too often been framed in a negative way. However an increasing body of work looking into individuals' resilience strategies demonstrates that users are not mere components of a system prone to fatigue, frailty and error. While it may be the case that in some cases, deviations in human performance can erode the resilience of a wider system, this does not negate the fact that frontline operators can and frequently do make a positive and largely underreported contribution to the resilience of a system. In displaying resilient qualities of their own, proactively recognising and mitigating potential risks and managing

threats to performance, users themselves represent a channel through which designers can enhance the resilience of system interactions.

Acknowledgements

We are grateful to *Dominic Furniss* and *Jonathan Back* for their extended guidance and input, and *Katarzyna Stawarz* and *Atish Rajkomar* for contributing strategy episode data. This work was supported by UK EPSRC grant [EP/G059063/1], *CHI+MED*.

References

[1] Bergström J., van Winsen, R., and Henriqson, E. On the rationale of resilience in the domain of safety: A literature review. Reliability Engineering and System Safety. 2015.

[2] Hollnagel, E. A tale of two safeties. Nuclear Safety and Simulation 4 (1), 1–9. 2013.

[3] Randell, R. and Johnson, C. User Adaptation of Medical Devices. In Proc. 21st European Conference on Human Decision Making and Control 105-109. 2002.

[4] Mumaw R, Roth E, Vicente K, Burns C. There is more to monitoring a nuclear power plant than meets the eye. In Human Factors; 42, (1) 36–55. 2000.

[5] Malakis S, Kontogiannis T. Cognitive strategies in emergency and abnormal training: implications for resilience in air traffic control. In Proc. The third Resilience Engineering Symposium. 2008.

[6] Furniss, D., Back, J., Blandford, A., Hildebrandt, M., & Broberg, H. A resilience markers framework for small teams. Reliability Engineering & System Safety, 96(1), 2–10. 2011.

[7] Furniss, D., Back, J., Blandford, A. Cognitive Resilience: Can We Use Twitter to Make Strategies More Tangible? Proc. ECCE '12. 96-99. 2012.

[8] Day, J., Buchanan, G., & Makri, S. Introducing the RSDiary App for the Collection of Resilience Strategies. Proc. ECCE '15. 2015.

Accounting for Organisational faults in Task Model Based Systematic Analysis of System Failures and Human Errors

Camille Fayollas, Célia Martinie, Philippe Palanque, Racim Fahssi
ICS-IRIT, University of Toulouse,
118 Route de Narbonne,
F-31062, Toulouse, France
name@irit.fr

Abstract. The overall dependability of an interactive system is the one of its weakest component which is usually its user interface. The presented approach integrates techniques from the dependable computing field and elements of user-centred design to provide a wider coverage of possible faults. Risk analysis and fault tolerance techniques are used in combination with task analysis and modelling to describe and analyse the impact of system faults on human activities and the impact of human deviation or errors on system performance and overall mission performance. A technique for systematic analysis of human errors, effects and criticality is proposed (HEECA). It is inspired and adapted from the Failure Mode, Effects and Criticality Analysis (FMECA) technique. The key points of the approach are: a) the HEECA technique combining a systematic analysis of the effects of system faults and of human errors, b) a task modelling notation to describe and to assess the impact of system faults and human errors on operators' activities and system performance. These key points are illustrated on an example extracted from a case study of the space domain. It demonstrates the feasibility of this approach as well as its benefits in terms of identifying opportunities for re-designing the system, re-designing the operations and for modifying operators' training. Lastly, a discussion presents the main challenges for also taking into account organisational faults in an integrated way with the proposed approach.

1 Introduction

The overall dependability of an interactive system is the one of its weakest component and there are many components in such systems ranging from the operator processing information and physically exploiting the hardware (input and output devices), interaction techniques, to the interactive application and possibly the underlying non interactive system being controlled. This paper proposes an approach integrating these aspects in order to address system and human dependability altogether. These two aspects of dependability are usually dealt with separately as the research contributions come from different

and usually unrelated scientific communities. In the dependable computing community, techniques have been proposed to cope with the impact of system failures and to assess it in a precise manner but operators' behaviour remains outside of the techniques. In the human reliability and in the human computer interaction communities, approaches have been proposed demonstrating the suitability of task modelling techniques to address system and human dependability analysis. This paper presents an integrated approach taking into account both system failures and human errors while designing interactive systems. This approach aims at leveraging existing techniques in the fields of: dependable computing, human reliability assessment and human computer interaction. The proposed technique also aims at providing complete and unambiguous task descriptions which support fine-grain analysis of both human and system aspects. Finally, this paper proposes to extend this approach in order to take into account organisational faults.

The article is structured as follows. Section II provides a brief review of types of system faults and human errors. Section III presents a task model-based stepwise process to describe and analyse the impact of system faults and human errors in an integrated manner. A case study from the space domain is then presented in section IV providing concrete application of the process presented in section III. Section V is dedicated to the identification of ways of integrating organisational faults inside the approach while section VI concludes the paper.

2 System faults and human errors

This section presents a review of types of system faults and human errors it provides the underlying information that is used in the process described in section III.

2.1 Considering system faults

2.1.1 Development faults

When considering the faults that can impair interactive computing systems, the first kind of faults that come in mind are the software development faults. These faults are introduced by a human (the developer) during the system development. They can be, for instance, bad designs or programming errors and they result in software defects [5] that can lead to software failures. In the area of software engineering, methods and techniques have been introduced to prevent such faults and include (formal methods, structured programming, OO

programing…). In the HCI community, a lot of work has been carried out for the prevention and removal of development software faults including software architectures (e.g. [2]), formal description techniques and verification (e.g. [17]), testing (e.g. [3]) or the use of debugging tools to remove human made faults.

2.1.2 Natural faults

In the domain of fault-tolerant systems, empirical studies have demonstrated (e.g. [18]) that software crashes may occur even though the development of the system has been extremely rigorous. One of the many sources of such crashes is called natural faults [1] triggered by alpha-particles from radioactive contaminants in the chips or neutron from cosmic radiation. A higher probability of occurrence of faults [27] concerns systems deployed in the high atmosphere (e.g. aircrafts) or in space (e.g. manned spacecraft [8]). Furthermore the evolution of modern IC components may lead in the next future to a higher probability of physical faults in operation. For instance, the recommendation for avionics systems is 100 FITs over 25 years lifetime, however, the current Deep Sub-Micron (DSP) technology may lead to a failure rate up to 1000 FITs, only during 5 years operational life time [25]. This is major worry in the avionics industry since this tendency has two bad sided effects, i) the reduction of the life time of the systems and ii) the increase of the failure rate due to hardware faults. Such natural faults demonstrate the need to go beyond classical fault avoidance at development time (usually brought by formal description techniques and properties verification) and to identify all the threats that can impair interactive systems.

2.2 Considering human errors

Several contributions in the human factors domain deal with studying internal human processes that may lead to actions that can be perceived as erroneous from an external view point. In the 1970s, Norman, Rasmussen and Reason have proposed theoretical frameworks to analyse human error. Norman, proposed a predictive model for errors [20], where the concept of "slip" is highlighted and causes of error are rooted in improper activation of patterns of action. Rasmussen proposes a model of human performance which distinguishes three levels: skills, rules and knowledge (SRK model) [21]. This model provides support for reasoning about possible human errors and has been used to classify error types. Reason [22] takes advantages of the contributions of Norman and Rasmussen, and distinguishes three main categories of errors:

1. Skill-based errors are related to the skill level of performance in SRK. These errors can be of one of the 2 following types: a) Slip, or routine error, which is defined as a mismatch between an intention and an action [20]; b) Lapse which is defined as a memory failure that prevents from executing an intended action.
2. Rule-based mistakes are related to the rule level of performance in SRK and are defined as the application of an inappropriate rule or procedure.
3. Knowledge-based errors are related to the knowledge level in SRK and are defined as an inappropriate usage of knowledge, or a lack of knowledge or corrupted knowledge preventing from correctly executing a task.

At the same time, Reason proposed a model of human performance called GEMS [22] (Generic Error Modelling System), which is also based on the SRK model and dedicated to the representation of human error mechanisms. GEMS is a conceptual framework that embeds a detailed description of the potential causes for each error types above. These causes are related to various models of human performance. For example, a perceptual confusion error in GEMS is related to the perceptual processor of the Human Processor model [4].

Causes of errors and their observation are different concepts that should be separated when analysing user errors. To do so, Hollnagel [9] proposed a terminology based on 2 main concepts: phenotype and genotype. The phenotype of an error is defined as the erroneous action that can be observed. The genotype of the error is defined as the characteristics of the operator that may contribute to the occurrence of an erroneous action.

These concepts and the classifications above provide support for reasoning about human errors and have been widely used to develop approaches to design and evaluate interactive systems [26]. As pointed out in [20] investigating the association between a phenotype and its potential genotypes is very difficult but is an important step in order to assess the error-proneness of an interactive system.

3 An integrated approach to account for system faults and human errors

This section presents the proposed approach to take into account for system failures and human errors at design and development time. It is composed of a stepwise process as well as modelling notation and tools [13].

3.1 A stepwise process to account for system faults and human error

The process for taking into account both system faults and human errors at design and development time is illustrated in Figure 3. The proposed process is decomposed in 7 phases:

1. Task analysis and modelling (similar to steps 1 and 2 of the FMECA analysis process).
2. Filtering out tasks and actions depending on the type of analysis to be performed
3. Effects and criticality analysis for human errors and system failure modes (similar to steps 3-5 of the FMECA analysis process).
4. Inventory of the couples {activity node, criticality} and inventory of the additional tasks that would be needed to recover from system failures and/or human errors (which matches step 6 of the FMECA analysis process).
5. Construction of enriched task models (models integrating potential system failures and human errors as well as articulatory tasks to recover from them).
6. Analysis of the impact of the system faults and human errors on the users' performance and on the global mission (system and organization).
7. Identification of design alternatives and proposals for modifying users' tasks and/or system's functions (which matches steps 6 of the FMECA analysis process).

3.2 Task modelling with HAMSTERS

HAMSTERS (Human – centered Assessment and Modeling to Support Task Engineering for Resilient Systems) is a tool-supported graphical task modelling notation for representing human activities in a hierarchical and structured way.

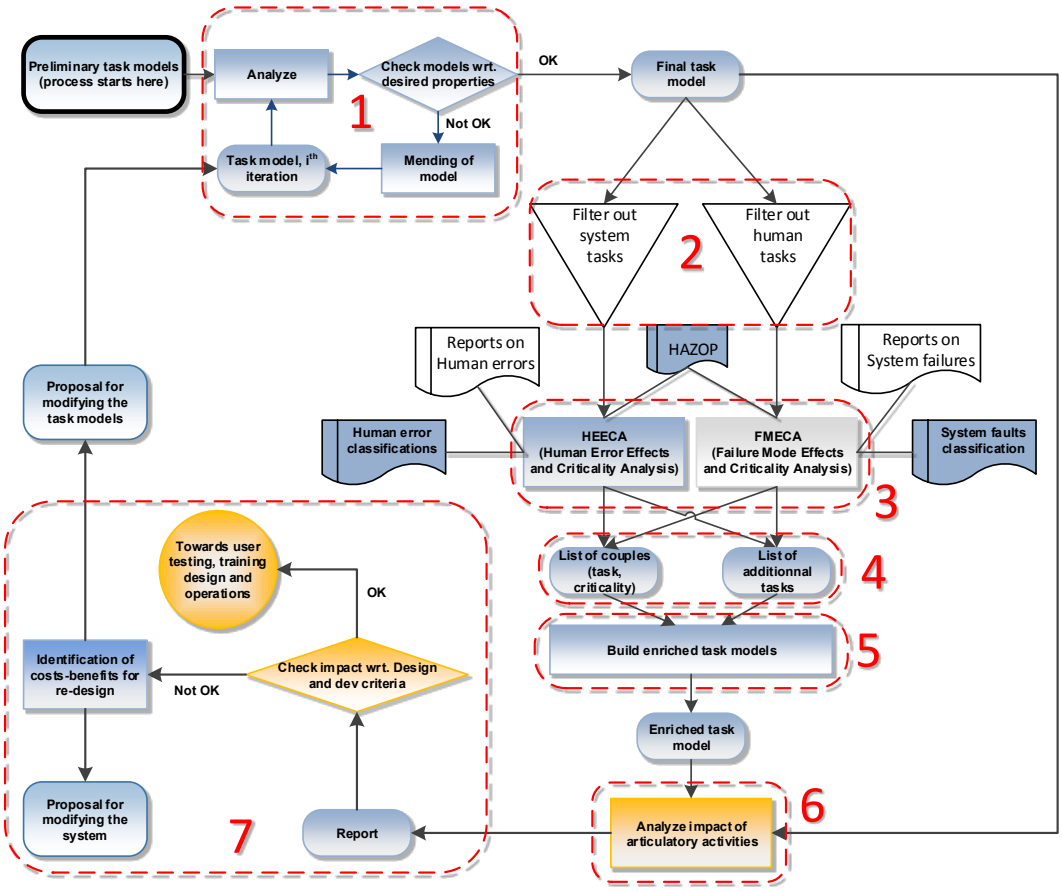


Figure 2. Process to account for system faults and human error during the design and development of an interactive critical system.

3.2.1 HAMSTERS notation

At the higher abstraction level, goals can be decomposed into sub-goals, which can in turn be decomposed into activities. Output of this decomposition is a graphical tree of nodes that can be tasks or temporal operators. Tasks can be of several types (cognitive, interactive, abstract…) and contain information such as a name, information details, and criticality level. Only the single user high-level task types are presented here but they can be further refined. For instance the cognitive tasks can be refined in Analysis and Decision tasks [12] and collaborative activities can be refined in several task types [10].

Temporal operators (based on LOTOS) are used to represent temporal relationships between sub-goals and between activities.

Tasks can also be tagged by properties to indicate whether or not they are iterative, optional or both. The HAMSTERS notation is supported by a CASE tool for edition and simulation of models. This tool supported notation also provides support for structuring a large number and complex set of tasks introducing the mechanism of subroutines [11], sub-models and components [7]. Such structuring mechanisms allow describing large and complex

activities by means of task models. These structuring mechanisms enables the breakdown of a task model in several ones that can be reused in the same or different task models.

HAMSTERS expressive power goes beyond most other task modelling notations particularly by providing detailed means for describing data that is required and manipulated [15] in order to accomplish tasks.

3.2.2 Relationship between HAMSTERS notation elements and genotypes

All of the above notation elements are required to be able to systematically identify and represent human errors within task models. Indeed, some genotypes (i.e. causes of human errors) can only occur with a specific type of task or with a specific element in a task model described using HAMSTERS. This relationship between classification of genotypes in human error models and task modelling elements is not trivial. For this reason, Table 2 presents the correspondences between HAMSTERS notation elements and error genotypes from the GEMS classification [117]. Such a correspondence is very useful for identifying potential genotypes on an extant task model.

It is important to note that strategic and situational knowledge elements are not present in this table. Indeed, such constructs are similar to the M (Methods) in GOMS and thus correspond to different ways of reaching a goal. As all the methods allow users to reach the goal an error cannot be made at that level and is thus not connected to a genotype.

Table 2. Correspondence between HAMSTERS elements and genotypes from GEMS [117]

Element of notation in HAMSTERS	**Related genotype from GEMS 117117**	
Perceptive task	Perceptual confusion (Skill Based Error) Interference error (Skill Based Error)	
Input task Motor task	Interference error (Skill Based Error) Double capture slip (Skill Based Error) Omissions following interruptions (Skill Based Error)	
Cognitive task	Skill based errors	Double capture slip Omissions following interruptions Reduced intentionality Interference error Over-attention errors
	Rule based mistakes	Misapplication of good rules First exceptions Countersigns and non-signs Informational overload Rule strength General rules Redundancy Rigidity Application of bad rules Encoding deficiencies Action deficiencies
	Knowledge based mistakes	Selectivity Workspace limitations Out of sight out of mind Confirmation bias Overconfidence Biased reviewing Illusory correlation Halo effects Problems with causality Problems with complexity
Information **Inf :** Information	Double capture slip, Omissions following interruptions, Interference error, all of the Rule Based Mistakes and Knowledge Based Mistakes	
Declarative knowledge **DK :** Declarative	All of the Knowledge Based Mistakes	

New notation elements, based on these correspondences, have been introduced to provide support for identifying and describing human errors in HAMSTERS task models [6].

4 Illustrative example: extract from the Picard Satellite Case Study

The Picard satellite dedicated to solar observation was launched by CNES in June 2010. We use a subset of it for our case study. This section presents an extract from a case study where the integrated approach has been applied to monitoring and control tasks performed with the Picard ground segment applications.

4.1 PICARD satellite ground segment

Satellites and spacecraft are monitored and controlled via ground segment applications in control centres with which satellite operators implement operational procedures. A procedure contains instructions such as sending telecommands (TC), checking telemetry (TM), waiting, providing required values for parameters.

Amongst the various ground segment applications used to manage the satellite platform, we focus on the ones that are used by controllers to ensure that the platform is functional. The platform has to be functional so that the mission (for which the satellite has been designed and developed) can be completed.

4.2 Controller's tasks analysis and modelling

Controllers are in charge of two main activities: observing periodically (i.e. monitoring) the vital parameters of the satellite and performing maintenance operations when a failure occurs. Depending on the satellite between thousands and tens of thousands parameters have to be monitored. The more frequent and relevant monitoring activities include observing: satellite mode, telemetry (measures coming from the satellite), sun array drivers statuses, error parameters for the platform, error parameters for the mission, power voltage (energy for the satellite), ground station communication status, and on board computer main parameters.

The "Start procedure" subroutine is presented in Figure 3. Fine grain modelling of users' actions with an interactive system is bound to the interactive system interface. The task models are highly dependent on the way the information is presented and reachable in the user interface. In this case study, the software application used by controllers is a procedure manager. The controller can select a procedure from the list and then s/he can start the procedure by pressing the "Start Procedure" button.

The procedure ("Search for procedure" iterative task). Once the controller has decided to select the procedure, the search task will be disabled (operator "[>") and the next task will be to:

- `Select the procedure ("Mouse selection [select procedure]" task, of subroutine type, in Figure 3)`
- `Start the procedure ("Mouse selection [start procedure] task, of subroutine type, in Figure 3)`

Finally, the system will start executing the procedure (system task in Figure 3). This task model also describes which information is required to reach the goal of starting a procedure.

The information about procedure reference. This information is required to be able to search for it in the list and to analyse that the targeted procedure has been found in the list (Box "I: (user) procedure reference with incoming and outgoing arrows to "Search for procedure" user task, "Perceive procedure to select" perceptive task and "Analyze that procedure is found" cognitive analysis task).

The information about the item (of the list) to be selected. Once the controller has decided to select the procedure, s/he produces new information which is the information about the item to be selected (Box "I: (user) item to be selected" with incoming arrow from the "Decide to select procedure" cognitive decision task and with an outgoing arrow to the "Mouse selection [select procedure]" subroutine task). Figure 4 presents the "Mouse selection" task model. It describes the fine grain actions that have to be performed for

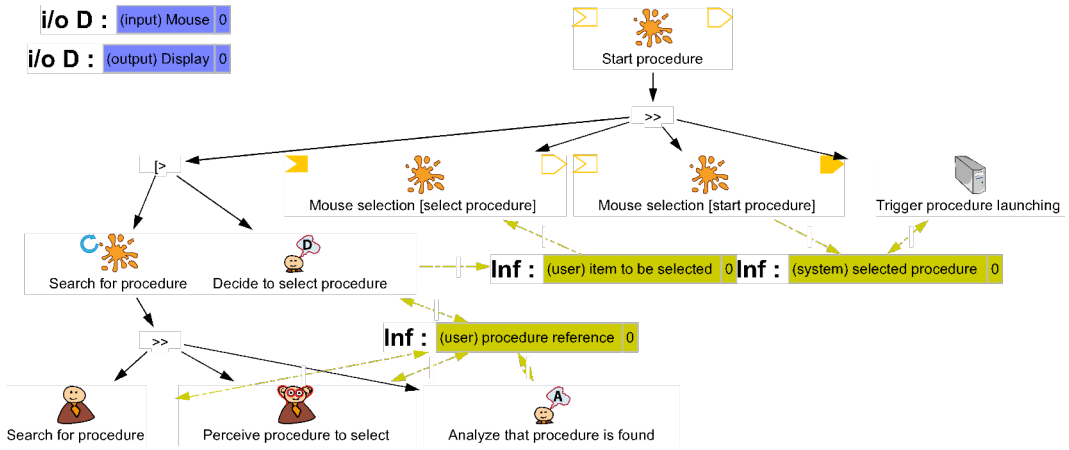


Figure 3. Task model of "Start procedure" task

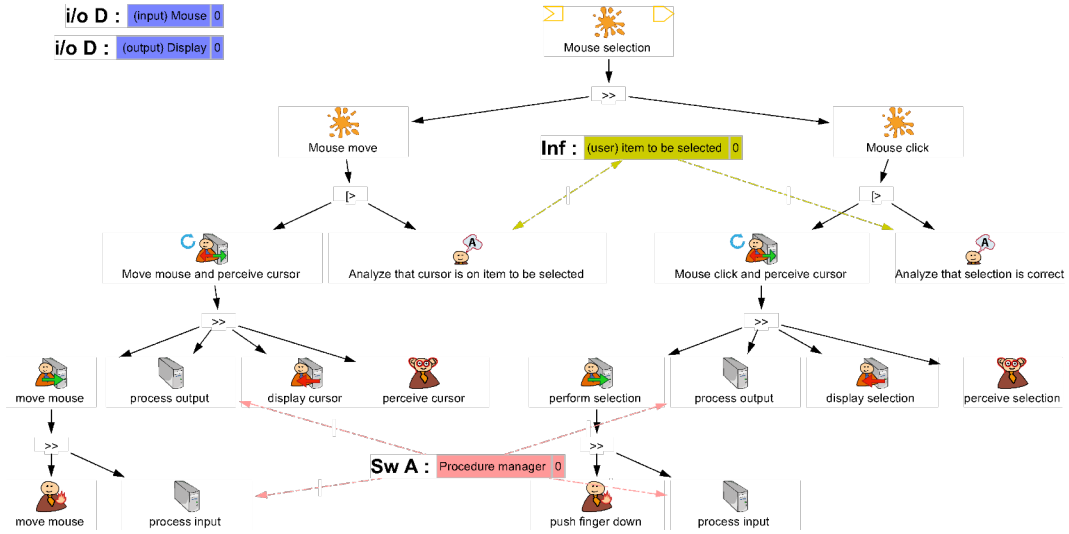


Figure 4. Task model of "Mouse selection" task

selecting a graphical object with a mouse device and pointer. It also describes the required information to reach this goal.

4.3 Human Errors, Effects and Criticality Analysis for the task of procedure selection, triggering and monitoring

Filtering out human actions from task models enables picking out the tasks and actions for which deviations and/or human errors may happen. The HEECA technique is then applied on these identified tasks and actions in order to systematically go through the potential issues and find out their criticality. Figure 6 contains an extract from the HEECA table for the controller's task of driving the execution of a procedure. For the rest of the example, we focus on the potential error related in line 3 of this table (surrounded with a bold rectangle). In this line, a critical issue is pointed out and would be caused by a perceptual confusion error when selecting the procedure to be launched. This error is related to the declarative information about the item to be selected that the controller has in mind (as depicted in Figure 5). S/he may analyse that the good item in the list has been selected whereas it is not. As described by the scenario, procedures can have names that differ only by a few characters, which may cause perceptual confusion errors.

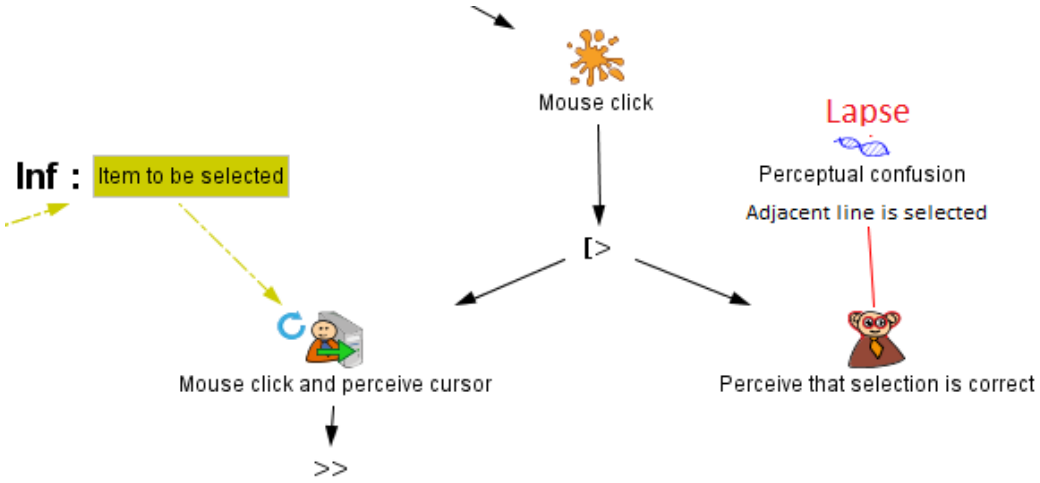


Figure 5. Focus on the action node where an error may occur

In this example, the task "Mouse selection: select procedure" may have several criticality levels depending on the identified scenarios and may reach the highest criticality levels. An erroneous or deviated behaviour during the mouse selection task may lead to delay in the mission and be tagged as critical. Figure 6 contains an extract from the HEECA table for the controller's task of driving.

#	Role	Scenario	Task node	Action node	Deviation (HAZOP)	Human error reference classifications	Error related to procedural or declarative information/ knowledge	Local effect	Effect on goal	Effect on mission	Severity	Proba	Criticality
1	Controller	The operator selects the procedure to trigger, but before the start it, he gets a call that made him start the procedure too late wrt. the satellite mode.	Start procedure SWITCH ON SADA2	Mouse selection: [start procedure]	Late	Not Applicable	Not applicable	Activity delayed	Task delayed	Delayed	Critical(3)	Remote (2)	6
2	Controller	The operator selects the procedure that is next to the good one, and he doesn't notice because he doesn't look the selected procedure on the display.	Mouse selection: [select procedure]	Perceive selection	No or not	Selectivity	Procedural	Another procedure is selected	Task interrupted	Delayed	Critical(3)	Remote (2)	6
3	Controller	The operator selects the procedure that is below the good one, and he doesn't notice because the both procedures have the same name by a few letters.	Mouse selection: [select procedure]	Analyze that selection is correct	Other than	Perceptual confusion	Declarative	Another procedure is selected	Task interrupted	Delayed	Critical(3)	Remote (2)	6
4	Controller	Operator select the wrong procedure but when he intends to change, he receives a call. After the call, the operator does not change the selection and starts the procedure.	Mouse selection: [select procedure]	Analyze that cursor is on item to be selected	Other than	Omissions following interruptions	Procedural	Another procedure is selected	Task interrupted	Delayed	Critical(3)	Remote (2)	6
5	Controller	The operator selects a procedure in a list, and before starting the procedure, he touches the "down" button without noticing, and the selected procedure is not the targeted one.	Start procedure SWITCH ON SADA2	Push finger down	Other than	Slip	Procedural	Another procedure is selected	Task interrupted	Delayed	Critical(3)	Remote (2)	6
...	...	...	...	...	...	...	...	...	...	...	...	...	...
20	Controller	Operator select the wrong procedure but when he intends to change, he receives a call. After the call, the operator does not change the selection and starts the procedure. The unintentionnally selected procedure causes the mission to fail.	Mouse selection: [select procedure]	Analyze that cursor is on item to be selected	Other than	Omissions following interruptions	Procedural	Another procedure is selected	Task interrupted	Mission fails.	Catastrophic (4)	Remote (2)	8
21	Controller	The operator selects a procedure in a list, and before starting the procedure, he touches the "down" button without noticing, and the selected procedure is not the targeted one. The unintentionnally selected procedure causes the mission to fail.	Start procedure SWITCH ON SADA2	Push finger down	Other than	Slip	Procedural	Another procedure is selected	Task interrupted	Mission fails.	Catastrophic (4)	Remote (2)	8

Figure 6. Extract from the HEECA table

5 Integrating the analysis of organisational faults

Organisational faults usually don't produce immediate faults (i.e. faults that trigger immediate failures and thus interruption of service) but introduce latent faults that under certain conditions and in presence of other faults (either being from the human or system side) [23] may cause failures. They belong to the main source of upstream factors to system (in its broader sense) failures that are leading to incidents and accidents.

STAMP (which stands for Systems Theory Accident Modelling and Process) has been developed by Nancy Leveson [16] and aims at providing a generic framework targeting (among other things) at identifying organisational faults.

One of the main aspects of STAMP is related to the notion of control (see Figure 7) and more precisely in the fact that control is not only based on human and system activities but also on the representations that are constructed by these two actors. For instance the box called Model of Automation inside the Human Supervisor bigger box represents the fact that the behaviour of the operator will be based on the mental representation that he/she owns about the behaviour of the automation currently deployed in the system he/she is supervising. This can explain many sources of accidents/incidents if the constructed mental model is different from the actual behaviour of the system. Such considerations are very similar to the ones that have been driving the field of Human-Computer Interaction over the last 30 years with the action theory from Donald Norman [19]. On the design side of physical or virtual objects the absence (or limited) of discrepancy between these models is named affordance.

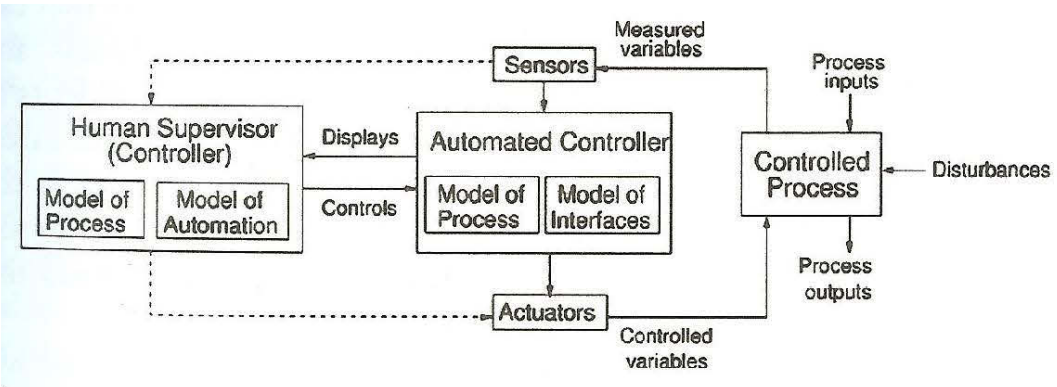


Figure 7. Human and automation roles in control (from [16])

Figure 8 highlights another important aspect of STAMP which is the necessity to take into account both design time and operation time when dealing with socio-technical systems. The detailed process also identifies

places where organizational constraints appear and how they can impact system safety and reliability both at design and operation time. STAMP is able to address at a high-level of abstraction a Large Scale Socio Technical System including organization aspects.

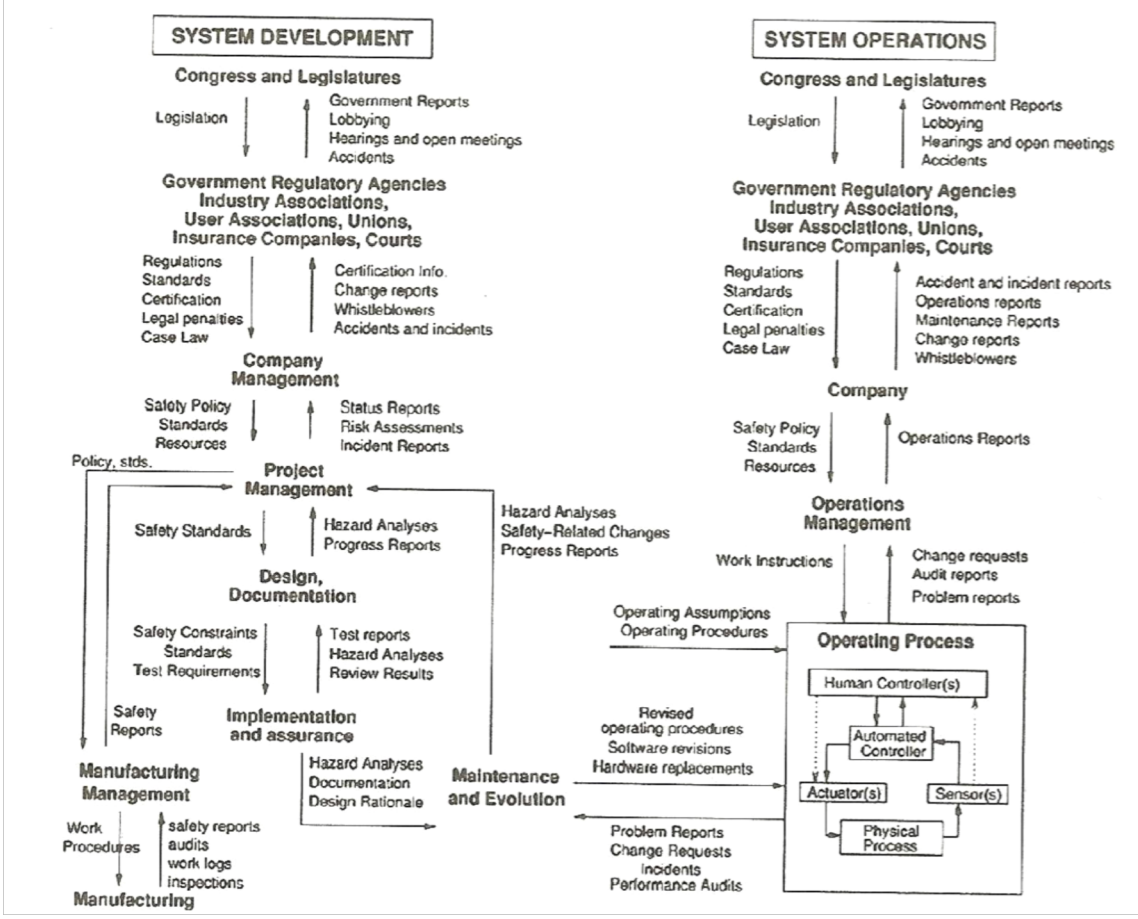


Figure 8. Integrated process dealing with System, Human and Organization views (from [16])

It is important to note that the STAMP analysis only remains at a very high level of abstraction, abstracting away from the details where actual system or human faults occur.

This high-level view will spread throughout the system design as guidelines or choices at design time. For instance the organization might decide to go for quick and cheap development processes that will end up with higher fault rates (especially development faults). Training programs will also influence occurrence of human errors and of operators' capabilities to deal with unexpected infrequent situations requiring a deep understanding of procedures and systems' behaviours that can only be acquired through experience and deep learning.

The control loop explained in the STAMP example provides another perspective based on information flow in the organization that can be made explicit through dedicated techniques such as workflow modelling and analysis [28]. Similar issues arise when modelling multiuser activities (even using HAMSTERS notation). This is where analysis of models can take place to identify (possibly following a high-level approach such as STAMP) the possible missing control and feedback loops.

6 Conclusion

This position paper presented an approach integrating techniques from dependable computing and user-centered design in order to improve the reliability of interactive systems. Risk analysis and fault- tolerance techniques are used in combination with task analysis and modeling to describe and analyze the impact of system faults on human activities and the impact of human deviation or errors on system performance and more generally on mission performance. A technique for systematic analysis of human errors, effects, and criticality is proposed (HEECA). It is inspired and adapted from the FMECA technique.

The key points of the proposed approach are: a) the HEECA technique combining a systematic analysis of the effects of system faults and of human errors, b) a task modelling notation to describe and to assess the impact of system faults and human errors on operators' activities and system performance. These key points have been illustrated on an example extracted from a case study of the space domain that has demonstrated the feasibility of this approach as well as its benefits in terms of identifying opportunities for re-designing the system, re-designing the operations and for modifying operators' training.

Finally, this paper discussed about the main challenges for integrating the analysis of organisation faults in the proposed approach.

References

[1] Avizienis, A., Laprie, J.-C., Randell, B., Landwehr, C. Basic concepts and taxonomy of dependable and se-cure computing. In IEEE Trans. on Dependable and Se-cure Computing, vol.1, no.1, pp. 11- 33, Jan.-March 2004.

[2] Bass, L., et al. (1991). The arch model: Seeheim revisited. In User Interface Developpers' Workshop.

[3] Bowen, J., & Reeves, S. UI-driven test-first development of interactive systems. In EICS 2011, ACM (2011), 165-174.
[4] Card, S., Moran, T., Newell, A. The model human pro-cessor: An engineering model of human performance. John Wiley & Sons, 1986.
[5] Chillarege, R.; Bhandari, I.S.; Chaar, J.K.; Halliday, M.J.; Moebus, D.S.; Ray, B.K.; Wong, M.-Y., "Orthogonal defect classification-a concept for in-process measurements," Software Engineering, IEEE Transactions on , vol.18, no.11, pp.943,956, Nov 1992
[6] Fahssi, R., Martinie, C., Palanque, P. Enhanced Task Modelling for Systematic Identification and Explicit Representation of Human Errors. In Proc. of IFIP TC 13 Intl. Conf. on HCI, INTERACT 2015, Bamberg.
[7] Forbrig, P., Martinie, C., Palanque, P., Winckler, M., Fahssi,R. Rapid Task-Models Development Using Sub-models, Sub-routines and Generic Components. Proc.of HCSE 2014, pp. 144-163.
[8] Hecht H. and Fiorentino E. Reliability assessment of spacecraft electronics. In Annual Reliability and Maintainability Symp., pages 341–346. IEEE, 1987.
[9] Hollnagel, E. Cognitive reliability and error analysis method (CREAM). Elsevier, 1998.
[10] Martinie, C., Barboni, E., Navarre, D., Palanque, P., Fahssi, R., Poupart, E., Cubero-Castan, E. Multi-models-based engineering of collaborative systems: application to collision avoidance operations for spacecraft. Proc. of EICS 2014, pp. 85-94.
[11] Martinie, C., Palanque, P., & Winckler, M. (2011). Structuring and composition mechanisms to address scalability issues in task models. In Human-Computer Interaction–INTERACT 2011 (pp. 589-609). Springer Berlin Heidelberg.
[12] Martinie, C., Palanque, P., Barboni, E., & Ragosta, M. (2011, October). Task-model based assessment of automation levels: application to space ground segments. In Systems, Man, and Cybernetics (SMC), 2011 IEEE International Conference on (pp. 3267-3273). IEEE.
[13] Martinie, C., Palanque, P., Fahssi, R., Blanquart, J.-P., Fayollas, C., Seguin, C. Task Model-Based Systematic Analysis of Both System Failures and Human Errors. IEEE Transactions on Human-Machine Systems, to appear in 2015.
[14] Martinie, C., Palanque, P., Fahssi, R., Blanquart, J-P., Fayollas, C., Seguin, C. Task Models Based Systematic Analysis of both System Failures and Human Errors. IEEE Transactions on Human Machine Systems, special issue on Systematic Approaches to Human-Machine Interface: Improving Resilience, Robustness, and Stability, to appear.
[15] Martinie, C., Palanque, P., Ragosta, M., & Fahssi, R. (2013, August). Extending procedural task models by systematic explicit integration of objects, knowledge and information. In Proceedings of the 31st European Conference on Cognitive Ergonomics (p. 23). ACM.

[16] Leveson N. A Systems-Theoretic Approach to Safety in Software-Intensive Systems. IEEE Trans. Dependable Sec. Comput. 1(1): 66-86 (2004).
[17] Navarre, D., Palanque, P., Ladry, J. F., & Barboni, E. (2009). ICOs: A model-based user interface description technique dedicated to interactive systems addressing usability, reliability and scalability. ACM TOCHI, 16(4), 18.
[18] Nicolescu B., Peronnard P., Velazco R., and Savaria Y. Efficiency of Transient Bit-Flips Detection by Software Means: A Complete Study. Proc. of the 18th IEEE Int. Symp. on Defect and Fault Tolerance in VLSI Systems (DFT '03). IEEE Computer Society, 377-384.
[19] Norman, D. The Design of Everyday Things (Originally published: The psychology of everyday things). New York: Basic Books, 1988.
[20] Norman, D. A. (1981). Categorization of action slips. Psychological review, 88(1), 1.
[21] Rasmussen, J. Skills, rules, knowledge: signals, signs and symbols and other distinctions in human performance models, IEEE transactions: Systems, Man &Cybernetics, 1983.
[22] Reason J. Generic error modelling system: a cognitive framework for locating common human error forms. New technology and human error, 63, 86. 1987.
[23] Reason J. Managing the Risks of Organizational Accidents. Ashgate Publishing Limited, 1997.
[24] Reason, J. Human Error, Cambridge University Press. 1990
[25] Regis, D.; Hubert, G.; Bayle, F.; Gatti, M., "IC components reliability concerns for avionics end-users," Digital Avionics Systems Conference IEEE/AIAA 32nd pp.2C2-1,2C2-9, 5-10 Oct. 2013.
[26] Ruksenas R., Curzon P., Blandford A., Back J. Combining Human Error Verification and Timing Analysis. EHCI/DS-VIS 2007: 18-35
[27] Schroeder B., Pinheiro E., and Weber W.-D.. DRAM errors in the wild: a large-scale field study. In ACM SIGMETRICS, pages 193–204, Seattle, WA, June 2009.
[28] van der Aalst W., ter Hofstede A. YAWL: yet another workflow language. Inf. Syst. 30(4): 245-275 (2005)

Modeling Monitoring Behavior for HMI Designs is Easy with the Right Tool

Bertram Wortelen[1], Sebastian Feuerstack[2], Marcus Behrendt[2]
[1]Carl von Ossietzky University, 26129 Oldenburg
bertram.wortelen1@uni-oldenburg.de
[2]OFFIS – Institute for Information Technology, 26121 Oldenburg, Germany
feuerstack@offis.de, marcus.behrendt.86@gmail.com

Abstract. Monitoring behavior is usually measured by performing studies in that eye-movements of operators monitoring a real system are recorded. Modern eye-tracking systems generate heat maps from recorded data. Areas of interest and the corresponding attention distribution can then be derived from the heat maps. We argue that recent progress in research indicates that psychological and physiological plausible human behavior models can often substitute time-consuming eye-tracking studies. Further on, with prediction models and the right tooling, heat maps can already be generated in early design phases: based on design sketches instead of running prototypes. We present a case study in the maritime domain in that experts predicted and analyzed their monitoring behavior. We argue that with the right tooling even non-experts will be able to predict their monitoring behavior. Enabling easy access to monitoring behavior prediction for everyone will improve future HCI design.

1 Introduction

Supervision and monitoring of complex systems is one of the main activities of an operator in a safety-critical environment such as air traffic control rooms, air plane cockpits, or ship bridges. More and more systems like cars, our office and home environments are getting smarter and act autonomously on behalf of us. This changes our role from being in active control to being there to observe and understand what is going on.

This has an impact on interface design. Being aware of how a design change might affect the human monitoring behavior is important information. Today the human monitoring behavior is often analyzed by recording eye movements, which requires observing humans controlling a working prototype of an interface design. We argue that psychological and physiological plausible human behavior models can often substitute time-consuming eye-tracking studies, can be earlier performed based on design sketches, and, if supported by a tool can also be performed by non-experts. In fact, analyzing

and predicting human monitoring behavior has been a research topic for several decades. In the mid of the 20th century a series of studies were conducted to analyze monitoring behavior of fighter jet pilots [3] in order to design cockpits with an optimized layout.

Since then, several models have been developed to predict human monitoring behavior in workplace environments like control rooms, cockpits, etc. In these environments the human behavior is strongly shaped by the tasks of the operator. A common prediction model assumption is, that humans try to perform tasks in an optimal way [6, 2, 8, 7, 5, 10, 1, 13]. Nearly all of these models argue that optimal monitoring behavior is based on knowledge of the probability distribution of information events for each information source and knowledge about the value of perceiving information events or respectively the costs of missing events. Wickens et al. [10] describe this as the two knowledge-based forces expectancy and value (of information events) that affect attention distribution. Besides these forces there are extrinsic factors like salient information events and the effort associated with sampling an information source that can lead to a potentially negative deviation from optimal monitoring [10, 9]. However, with sufficient training the knowledge-based factors should explain the majority of variance in monitoring behavior [5, 11].

Even though there is a vast knowledge base on monitoring theory, the above mentioned models are typically only used by human factors experts, because in order to use any of the above mentioned models to predict optimal monitoring behavior, a quantification of expectancy and value coefficients for each information source is required. We believe that this is a barrier for a more wide-spread adoption of attention models in industrial practice and we believe that a suitable software tool can enable non-experts to do this task. Furthermore we believe that even if the modeling predictions are not very accurate the modeling process itself is useful for explicitly extracting the knowledge of experts in a structured way.

In this contribution we present a tool for the quantification of value and expectancy coefficients for a given set of information sources. These coefficients can directly be used for the value and expectancy factors of Wickens' SEEV [10] and A–SA [11] models or the recent implementations of the SEEV model in the cognitive architectures CASCaS [12] and MIDAS [4]. The next section describes a case study for monitoring behavior prediction in the maritime domain that was partially tool-supported. Based on the experiences gained, Section 3 proposes a complete tool-supported prediction process to facilitate the quantification of coefficients that is targeted also to non-experts. Section 4 summarizes our position.

2 Predicting Maritime Chart Monitoring Behavior

Three different design variations for a nautical chart display were evaluated to test the effect that offering a new kind of information has on operator's monitoring. Four experts with different backgrounds participated in this qualitative study [3]. All were able to understand and successfully perform the

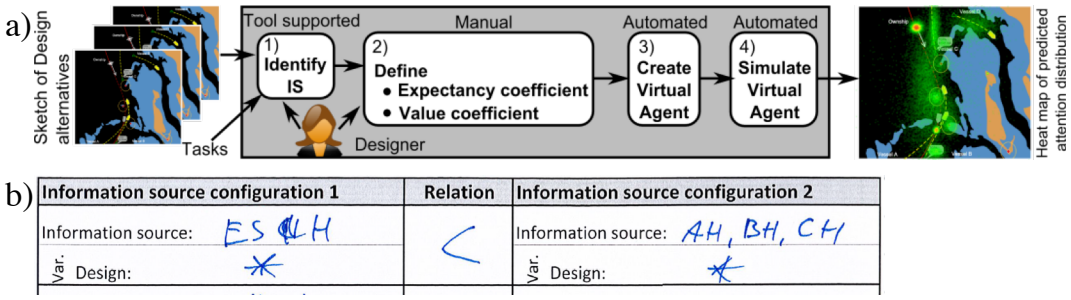


Fig. 1. (a) Tool-supported process to predict operator's attention allocation. It was executed by the HMI designer and the operator [3]. (b) Paper form to specify expectancy relation.

process of generating attention predictions. Fig. 1 illustrates the steps that the participants performed. In the first step, which already was tool-supported, they visually marked all sources of information (IS) on each design sketch. Different to what we expected, a high amount of IS were identified (between 18 and 47) and it figured out hard for them to remember all of them in the subsequent steps. We therefore supported them by a list with all IS identified. We also introduced abbreviations for each IS.

The second step (expectancy and value identification of all IS) was performed manually by filling paper forms (c.f. Fig. 1b). The expectancy of information events occurring in an IS was identified by mathematical relations ("<": greater, ">": smaller, and "=": identical) between IS that are relevant for a predefined set of tasks. For example the relation depicted in Fig. 1b states that IS "ESH" is expected to provide less new information than IS "AH", "BH", or "CH" (for all design variants). Since the same IS sometimes occurred on more than one design, with each IS also the corresponding design name was stated in the relation and we introduced a "*" operator, which could be used to refer to an IS on all designs. In a similar way a "*" operator can also be used to refer to all ISs on a specific design.

Unfortunately due to high number of ISs and complexity of the relation, one participant created an inconsistent relation that we were not aware of during the study. We could observe that all participants had trouble with the "mathematical" way of specifying the relation: It took them time to get comfortable with the notation and some required help to correctly write down what they intend to say. After a participant was finished with defining the

relation, we systematically walked though all IS identified and asked if this IS has or needs to be considered in the relation. This was done to ensure that the participant has at least considered all IS as candidates for including them in the relation. In a similar way, participants rated operator tasks by creating a task relation based on the task value. Additionally the relevance of an IS for a certain task was rated as either “necessary”, ”helpful”, or ”not relevant” on a form. All forms were then transcribed into an Excel table. Expectancy and value coefficients were calculated using the lowest ordinal algorithm [6, 15, 13], which derives the coefficients from the rank order of the IS in the expectancy relation, respectively the task in the task value relation.

The predictions were generated in a reasonable amount of time (between 42 and 160 minutes modeling time, without transcription time). Between 44% and 54% of this time was spent on identifying the expectancy relations, values of each IS, and by filling the paper forms. The next section details the tool-support and states how we considered the study observations for implementing a complete tool-supported attention prediction process.

3 Towards a structured, tool-driven process to compute attention predictions

We think that a complete tool-driven attention prediction process can resolve all issues that required our intervention in the maritime study, will reduce the overall modeling time, and also will enable non-experts to create attention predictions. The IS identification step was already tool supported in the study, thus we will focus on the tool-based generation of the expectancy and value coefficients. Based on the time measurements of the study the former one required most parts of the modeling time for all participants, whereas the value definition only took 8.2 minutes on average [3].

We decided for a two-step approach to define the expectancy relations: First, the user is asked to roughly build up a hierarchy of IS by dragging them from a list (Fig. 2c) into a hierarchy (Fig. 2b). Relations are automatically created while the hierarchy is established (Fig. 2d - for each two consecutive layers in the hierarchy a relation statement is created). In the middle of each relation statement there is a button indicating the current operator that can be clicked to toggle between “<”, ”>”, and ”=”. Second, new statements can be added and existing ones can be updated by directly editing the relation list of relation statements (Fig. 2d). The list of all IS (which is the source for dragging out “IS”) can be sorted alphanumerical, per design (like shown in Fig. 2c) and most importantly can be sorted by “usage” to ease identifying IS

that have not been defined in a relation statement so far. Inconsistent relation statements are instantly highlighted red (e.g. the last one in Fig. 2d).

We observed during the study that the participants intuitively grouped similar IS logically to reduce the writing effort (e.g. all "lighthouses" or all "shoals" in the study). Therefore we introduced IS grouping support to the tooling: The selection and dragging of several IS from the IS list (Fig. 2c) ends up in a popup window (Fig. 2a) enabling a fine grained selection of the

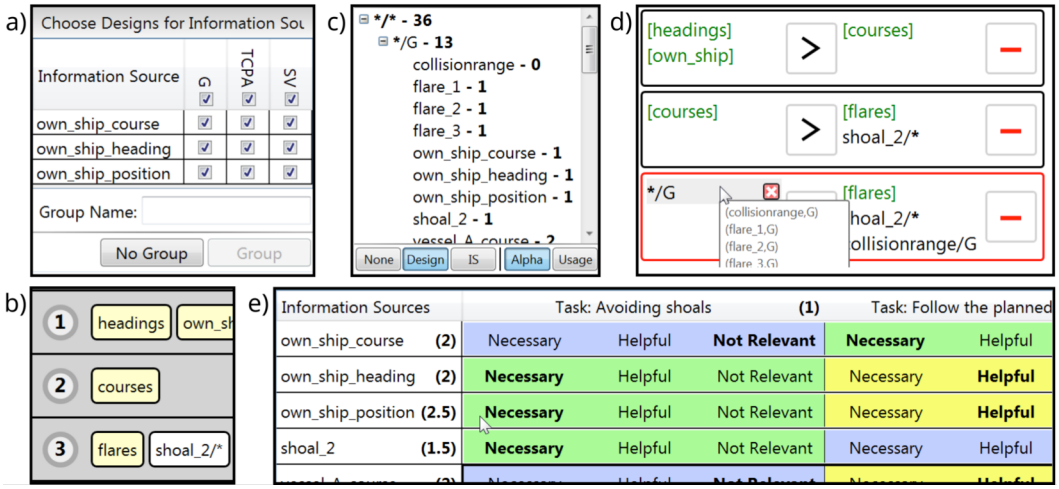


Fig. 2. Tool screenshots. (a) Popup window for defining an IS group. (b) Hierarchy of ISs. (c) List of all ISs. (d) Three expressions of an expectancvy relation. (e) Relevance matrix.

designs (columns) that should be part of a group. Additional, a group name can be defined (e.g. "own_ship" in Fig. 2a) that is then used in the relations (Fig. 2d – groups are identified by square brackets). Having the mouse pointer hovering above a group name or any other shortcut operator (like the "*" operator in Fig. 2d that sums up all IS from Design "G"), a tooltip shows all IS inside the group.

For the definition of the IS, we kept the idea of the paper form that used a relevance matrix to rate the value of each IS for a certain task with either "Necessary", "Helpful", or "Not Relevant" (Fig. 2e). Since the study participants were already fast with the rating process, we only added a coloring scheme to ease the identification of IS that have not been rated and also to support the user in perceiving, which tasks require a lot of IS or which IS is relevant for many tasks.

After the relations and relevance matrix have been defined the consistency of relations and matrix are automatically checked. Then the expectancy and value coefficients are calculated and fed into the following process step for creating the virtual agent. Manual transcription is no longer required. Up to this point the user worked solely with the mouse, except for naming the groups. The need for writing down the relation in the mathematical notation is

avoided. Initial relations are automatically created when setting up the expectancy hierarchy (Fig. 2b). We expect that this eases the understanding of the basic structure of the expectancy relation. The manual addition of new relations is also based on drag-and-drop operations between the IS list (or the hierarchy view) and the fields of the relations to further reduce typing effort.

4 Conclusion

We presented a tool that supports all steps required to characterize optimal monitoring behavior. All actions that the users have to do are kept as simple as possible. In fact, besides the domain knowledge of the monitoring task we expect that very little knowledge about the underlying monitoring model is required. The calculation of expectancy and value coefficients is automatically done by the tool. The required domain knowledge is extracted by guiding the user through a series of actions. Each statement of the relation (c.f. Fig. 2d) or entry in the relevance matrix captures a single aspect of the domain knowledge, e.g., “information events occur more frequently in information source A than in information source B” or “information source C is helpful for performing task D”. The integration of all these statements in an overall model and the calculation of the resulting coefficients is entirely performed within the tool.

In a follow-up action, we will conduct a study in the automotive domain to test how valid the predictions of the users are. We will ask experienced drivers to characterize monitoring behavior in a set of different driving situation using our tool. We intend to only offer a short video tutorial to teach the participants the tool usage. In parallel a driving simulator study is conducted for all of these driving situations. The predicted attention distribution will be compared with drivers’ attention distribution measured with an eye tracker. Other studies [6, 15, 13] that used human factor experts and the lowest ordinal heuristic to calculate the value and expectancy coefficients reported correlation coefficients for percentage dwell times for all information sources in the range of $0.60 <= r <= 0.98$. These studies were performed on a similar level of abstraction in realistic driving and flight simulators. On average over all participants we expect to achieve a correlation within this range.

References

[1] Byrne, M.D., Kirlik, A.. Integrated modeling of cognition and the information environment: A multilevel investigation (process and product modeling) of attention allocation to visual displays. Technical Report AHFD-04-14/NASA-04-4, Aviation Human Factors Division (AHFD), Institute of Aviation, University of Illinois, Urbana-Champaign, Savoy, Illinois 61874, Sept. 2004.

[2] Carbonell, J.R.: A queuing model of many-instrument visual sampling. IEEE Transactions on Human Factors in Electronics, 7(4):157–164, Dec. 1966.

[3] Feuerstack, S., Wortelen, B.: Revealing differences in de-signers' and users' perspectives: A tool-supported process for visual attention prediction for designing HMIs for maritime monitoring tasks. In Proceedings of 15th IFIP TC.13 International Conference on Human-Computer Interaction - INTERACT, Bamberg, Germany, 14.-18. Sept. 2015.

[4] Fitts, P.M., Jones, R.E., Milton, J.L.: Eye movements of air-craft pilots during instrument-landing approaches. Aeronautical Engineering Review, 9(2):24–29, 1950.

[5] Gore, B.F., Hooey, B.L., Wickens, C.D., Scott-Nash, S.: A computational implementation of a human attention guiding mechanism in MIDAS v5. In Duffy, V.G. (eds) Digital Hu-man Modeling, vol. 5620/2009 of Lecture Notes in Computer Science, pp. 237–246. Springer, Berlin, 2009.

[6] Horrey, W.J., Wickens, C.D, Consalus, K.P.: Modeling drivers´ visual attention allocation while interacting with in-vehicle technologies. J. Exp. Psychol.-Appl., 12(2):67–78, 2006.

[7] Moray, N.: Monitoring behavior and supervisory control. In Boff, K.R., Kaufman, L., Thomas, J.P. (eds) Handbook of Perception and Human Performance. vol 2. Cognitive Processes and Performance, chapter 40, pp. 40–1 – 40–51. Wiley-Interscience, Nov.1986.

[8] Senders, J.W.: The human operator as a monitor and controller of multidegree of freedom systems. IEEE Transactions on Human Factors in Electronics, HFE-5(1):2–5, Sept. 1964.

[9] Sheridan, T.B.: On how often the supervisor should sample. IEEE Transactions on Systems Science and Cybernetics, pp. 140–145, 5 1970.

[10] Smallwood, R.D.: Internal models and the human instrument monitor. IEEE Transactions on Human Factors in Electronics, HFE-8(3):181–187, 1967.

[11] Steelman-Allen, K.S.: Modeling the Control of Visual Attention in Complex Workspaces. PhD thesis, University of Illinois at Urbana-Champaign, Urbana, Illinois, 2011.

[12] Wickens, C.D., Helleberg, J., Goh, J., Xu, X., Horrey, W.J.: Pilot task management: Testing an attentional expected value model of visual scanning. Technical Report ARL-01-14/NASA-01-7, University of Illinois, Aviation Research Lab, Savoy, IL, Nov. 2001.

[13] Wickens, C.D, McCarley, J.S., Alexander, A.L., Thomas, L.C., Ambinder, M., Zheng, S.: Attention-situation awareness (A-SA) model of pilot error. In Foyle, D.C., Hooey, B.L. (eds) Human Performance Modeling in Aviation, pp. 213–239. CRC Press/Taylor & Fran-cis Group, New York, 2008.

[14] Wortelen, B.: Das Adaptive-Information-Expectancy-Modell zur Aufmerksamkeitssimulation eines kognitiven Fahrermodells. PhD thesis, University Oldenburg, 11 2014.

[15] Wortelen, B., Baumann, M., Lüdtke, A.: Dynamic simulation and prediction of drivers' attention distribution. Transportation Research Part F: Traffic Psychology and Behaviour, 21:278–294, 2013.

Workshop:
Human Work Interaction Design (HWID): Design for Challenging Work Environments

Interaction Design for Stratigraphic Analysis in Archaeology

Barbara Rita Barricelli, Stefano Valtolina
Department of Computer Science
Università degli Studi di Milano
Via Comelico 39/41, 20135 Milano, Italy
{barricelli, valtolin} @ di.unimi.it}

Abstract. Stratigraphic analysis, as defined my Harris in his Matrix approach, constitutes one of the most diffused analysis methods in archaeology. After a long collaboration with Archaeologists and other domain experts, we developed a second version of ArchMatrix, a system for graph-based interaction with excavation data. This system aims at solving the challenges that affect the archaeological practice given by its peculiar distributed environment. This position paper presents an overview of the research domain, the ArchMatrix system and the results of the preliminary usability tests with the end users.

1 Introduction

Archaeology is the study of human history by means of the analysis of signs left by mankind in the world. Everything that is buried underground tells us a story about how the human presence evolved over time. Archaeological excavations allow not only to find artefacts but also to understand different historical phases that succeeded in history leaving traces that can be studied through stratigraphic analysis. There are many common points between geology and archaeology because they both study the transformations of the soil during the years by observing the stratigraphies, i.e. by analyzing the sequence of soil layers. Such layers are originated by erosion and backlog phenomena that correspond respectively to creation and destruction of soil that have led to radical transformations of the environment. In geology, the various sequences of layers can be sorted in chronological order according to their vertical underground position: a layer lies on another more ancient layer and its rocks can be dated by analyzing them and assigning them to specific historical periods. Archaeology deals with stratigraphic analysis in a slightly different way: there is no analysis of rocks but of the different components, called stratigraphic units that can have human or natural origins. Such stratigraphic units can be seen as “positive”, if they are related to backlog activity (e.g., pavements, walls) or “negative” if they are originated by a destruction activity (e.g. excavations). The analysis of the relationships that

exist between the various units constitutes what is the archaeological stratigraphic analysis that was formalized by Edward C. Harris [1]. The archaeological domain is characterized by strong social and organizational factors and the successful development of our research in this area proves once more the validity of the Human Work Interaction Design framework [7] that strongly relies on the influence of environment and context on interaction in working places.

This paper presents the evolution of a research and design work performed in the last 3 years that we published in its early stage of development in [4]. We designed and developed ArchMatrix, an interactive system aimed at supporting archaeologists in archiving, managing and studying the findings collected during archaeological excavations. The design of the system followed a participatory action design research [6] and was based on a first prototype developed with technologies and languages that are today obsolete. The new system offers better performances and according to the results of the evaluations, solves all the issues highlighted during the evaluation of the first prototype.

2 Challenges in Interaction Design for Archaeological Practice

The process followed by archaeologists on excavation sites is made of seven main steps:

1) Detection of the most recent stratigraphic unit, that is the one that covers all the others and that is not covered by any other units.
2) Definition of the limits of the most recent stratigraphic unit and of its (physical) relationships with other units. This is one of the most important phases of excavation and is strictly related with the personal and subjective evaluation made by the archaeologists. This task needs to be done in a highly controlled environment to preserve its integrity and to avoid external contaminations.
3) Written, photographic, and graphical documentations of the characteristics of the soil before its exploration. Keeping track of the original status and successive modifications of the environment allows the archaeologists to analyse their findings even later their discovery and to preserve information about their original conditions.
4) Choice of the most suitable tools to be used during the exploration of the stratigraphic units to be removed. Large machines or small tools can be used.

5) The actual excavation of the soil.
6) Documentation of everything that has been detected, excavated, discovered. It is a highly challenging task because it needs a very deep precision by the archaeologists and everything that will be inserted in the post-excavation documentation will be used to interpret the findings and its quality will deeply influence the final results. The post-excavation documentation has to be focused on both vertical and horizontal plane of excavation.
7) Analysis of the collected data. Once the excavation site and its documentation are complete, this will be stored in a safe place and everything that has been found will be "indexed" by the documentation referring to the stratigraphic unit that it belonged to.

What makes this application domain very challenging is its distributed nature: many different tasks are performed in different places, by numerous teams of researchers and archaeologists and very often in multidisciplinary contexts involving experts in different domains (e.g., chemistry, geology, and historian). Moreover, excavation sites offer an uncomfortable work environment, most of the time very dusty and in some seasons and locations very sunny and hot. These specific environmental characteristics affect the use of mobile devices, especially for what concerns the visibility and legibility of screens. In the last years, we studied the archaeologists practice on the excavation sites to find out how to address the open problems and how to design the most suitable interaction and user experience for such a peculiar work environment.

In particular, we explored the graph metaphor as a visual strategy for enabling a fast and effective interaction. In archaeology, graph visualization systems have to face the problem of facilitating the exploration and analysis of a vast amount of data by means of visual methods and tools able to support needs of a wide array of different research communities involved in the study of an excavation such as archaeologists, architects, geologists, chemists, and biologists.

Information visualization strategies are applied in archaeology for assisting domain experts in the examination and interpretation of the stratigraphy of excavation sites, and identifying both natural and cultural strata. During the excavation of an archaeological site, when a stratigraphic unit is detected, the archaeologists fill in a proper form to keep track of its properties and the relationships among the unit and other units are made explicit. Three types of relationship could exist: active, passive, and neutral. Active relationships are a) covers, b) fills, c) leans, and d) cuts; passive relationships are a) is covered, b) is filled, c) relies on, d) is cut; neutral relationships are a) is equal to, and b)

binds. Once all the properties and relationships' data are collected in the form, the archaeologists represent the stratigraphic unit in a drawing on paper that will be used in successive data analysis. At the end of the excavation campaign, all this material is collected and studied for creating the correspondent Harris Matrix. The resultant Harris Matrix enables the archaeologists to determine the chronology of the various units (an example is shown in Figure 1). The rectangles represent the stratigraphic units, while lines are used to indicate the existing relationships among them (e.g. "copre" means "covers", "taglia" means "cuts", "riempie" means "fills"). After the drawing is complete, the archaeologists determine the historical age to which the stratigraphic units belong to. The Harris Matrix gives information about the temporal collocation of the units, while the analysis of the materials found in the unit allows to retrieve exact dates. The next activity is the detection of the overlays, i.e. logical levels that are constituted by the stratigraphic units that belong to the same historical period.

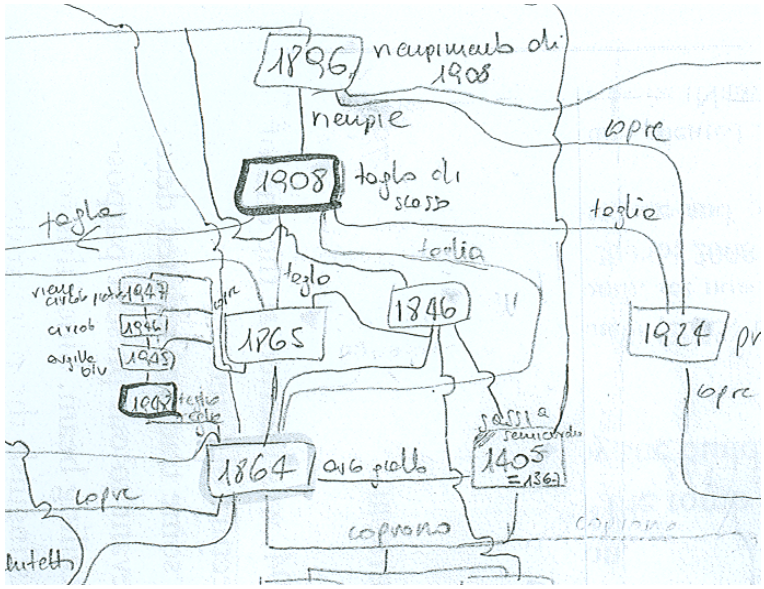


Figure 9. A section of a Harris Matrix diagram.

This work is usually done using paper-based documents, and the more the data grows, the more the consultation becomes challenging. Moreover, the more stratigraphic units are added, the more the Harris Matrix to be drawn becomes big and becomes extremely complicated to be modified and extended. Given these difficulties, the main disadvantages that arise from the use of a non-digital approach is the impossibility of properly diffuse the knowledge that is gathered through the archaeologists' activity performed on the field, the difficult of keeping the Harris Matrix documents up to date. In fact, to give permissions to modify the Matrix to more than one person is not

so simple, and new problems arise when many people have access and modification permission to the same resource. Visual interactive systems represent therefore the best choice to support archaeologists' activities and collaboration.

3 ArchMatrix Visual Interactive System

In 2012 we designed a visualization system named ArchMatrix with the collaboration of a team of archaeologists that was involved through the application of participatory design's principles. A set of evaluations on the system highlighted the presence of several usability problems that needed to be solved and its use in time pointed out the need of some changes to enable collaboration between the members of the teams.

Given the technological advancements happened in the last years and an evolution of the archaeologists and the other stakeholders in the management of data through a graph-based interaction, we decided to redesign the system by involving the experts again and to redevelop the system by using the most recent technologies and languages. A screenshot of one of the principal view of the ArchMatrix system is depicted in Figure 2.

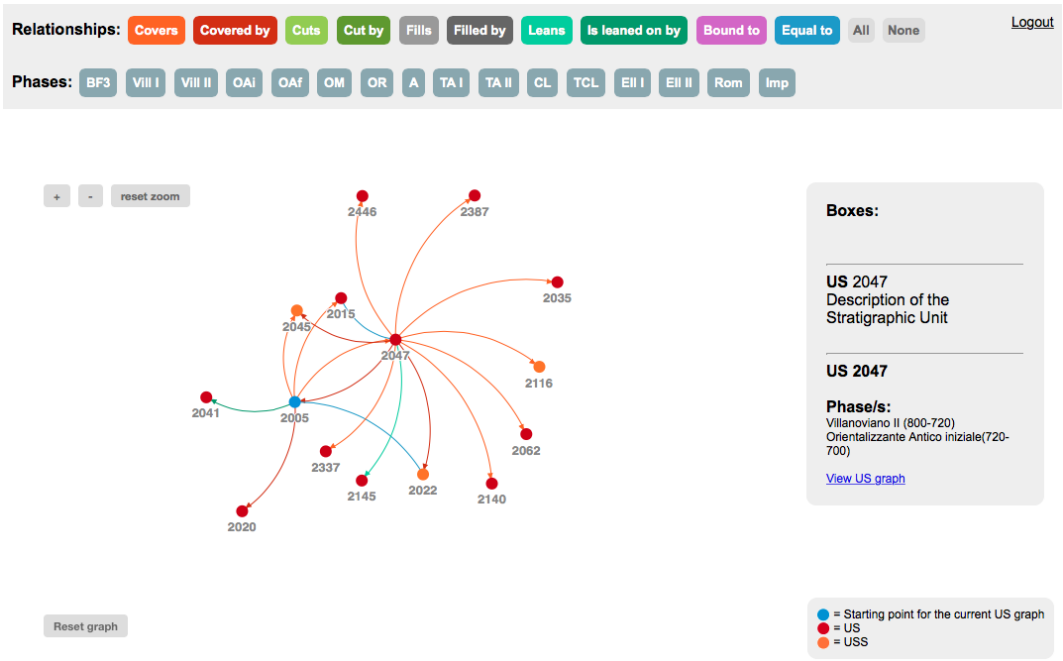


Figure 2. Harris Matrix visualization in the ArchMatrix system. The relationships of a selected stratigraphic unit are shown by means of a graph-based metaphor.

This new version of ArchMatrix allows its users to perform specific queries on the stratigraphic units data and to discover and analyse relationships

between them. An important feature that completes the digitalization of the Harris Matrix traditional method is the possibility of classifying the stratigraphic units according to specific time period. This allows the archaeologists to discuss with the other members of the team about the provenance and dating of findings.

It is through the discussion feature, implemented as an annotation tool, that ArchMatrix becomes a system for supporting collaborative analysis of archaeological excavation sites and related findings.

The graph-based interface gives the opportunity to update the information related to the stratigraphic units without having to deal with databases and complex textual interfaces.

From a technical point of view the graph provides researchers with a visual representation of the stratigraphic units highlighting geometric, topological and temporal relationships. Stratigraphic units are necessary to detect the relative chronological sequence of the entire excavation site but they also produce a number of supplemental data that are not included in the classic tool used for stratigraphic visualization. To support the complex and interdisciplinary decision-making activity at the base of the archaeologists' work, ArchMatrix can be used for visualizing and handling graphs through which researchers can define queries and algorithms able to explore stratigraphic units and the knowledge retrievable from excavation database or other knowledge sources such as data from geographical systems. This solution allows archaeologists to develop new opportunities for their investigation (both individual and collaborative), to increase their knowledge, to improve their traditional working practices and to develop new ones.

4 Evaluation

As a first preliminary evaluation we performed a usability user test followed by a questionnaire (SUS [2] and CSUQ [3] standards). The 4 participants were experts in the application of the Harris Matrix method and demonstrated to be able to use the system in a quite easy way. SUS questionnaire result was 83/100, while CSUQ reported a 4/5 on all the metrics (use of the system, quality of information, quality of interaction, and overall evaluation).

We tried to investigate about the problems that were highlighted during the evaluation of the first prototype of ArchMatrix and we noticed that in this second version all of them have been successfully solved. The problems that affected the first prototype were mainly related to navigation. In this version, the quality of the navigation approach has improved and the users demonstrated that is quite easy and fast to find all the information in this new

prototype. The annotation tool that allows to keep track of the motivations of changes made by the users has been very positively received and responds to the comments and requests gathered during the evaluation of the first prototype. An interesting outcome from the final unstructured interviews done in this evaluation, is that the archaeologists do not appreciate to see the term "graph" used in the system, even if they perfectly know that the visual structured used for the US/USS representation is in fact a graph: they find it confusing to use the term "graph" and the term "(Harris-)matrix" to intend the same thing.

The system is currently under testing by the archaeologists and we are collecting feedbacks and analysing research results (in terms of both qualitative and quantitative data). Only a long-time use in practice will point out any possible problems with the system.

5 Further works

Further studies aim at exploiting the ArchMatrix to support domain experts in defining simulations based on 3D reconstructions of archaeological evidence that no longer exists according to probabilistic hypotheses and inferred interpretations. By exploiting data of repositories containing 3D models of archaeological fragments and by integrating them with the representation of the stratigraphic units offered by ArchMatrix, this new service will provide a sophisticated recommendation system (RS) able to support researchers to carry out hypotheses simulations of 3D reconstructions. Traditional RS such as: Collaborative Filtering Systems, Content-Based Systems, and Hybrid Recommendation Systems [5], can be used for designing a dynamic support in terms of recommendation strategies that combine metadata representation of the distributed archaeological evidences, the domain experts' expectations, wishes and competences.

Specifically, this service will enable hypotheses simulations starting from analyses of 3D fragments and models collected according to similarity relationships about their semantics, context of provenance, chronology, geometrical compatibility or other metadata. A distinguishing recommendation approach that will be used concerns the use of learning algorithms to identify relations between the features of the integrated contents that may prove, as well as indications suitable to the inquirer. Thanks to the massive use of metadata, the archaeological simulation may be homogeneously identified through a vector of parameters (the metadata representation). In addition, thanks to the feedback of the previous users, a score can be associated with each composition of vectors in simulations. This

enables a dynamical decision tree procedure where, according to the current choice of the user, the service will propose branches of decision trees that may lead to satisfactory completion of the simulation, possibly listed in a monotone ranking. The results of this activity will be the creation of virtual reconstructions of original archaeological objects defined according to probabilistic hypotheses depending on the features of the combined fragments (e.g. chemical composition, chronology, semantics) and the comparison with other 3D models coming from different contexts.

Acknowledgements

The authors wish to thank the Archaeologists and other experts who collaborated with us in the design and evaluation of ArchMatrix and Alberto Faleschini for his work during his Master Degree Thesis.

References

[1] Harris, E. C. 1979. Principles of Archaeological Stratigraphy. Academic Press, New York, NY.

[2] Brooke, J. 1996. SUS: a "quick and dirty" usability scale. In *Usability Evaluation in Industry*, P. W. Jordan, B. Thomas, B. A. Weerdmeester, and A. L. McClelland, Eds. Taylor and Francis, London, UK, 189-194.

[3] Lewis, J. R. 1993. *IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use*. Technical Report 54.786. IBM.

[4] Barricelli, B.R., Valtolina, S., Marzullo, M. (2012). ArchMatrix: A Visual Interactive System for Graph-based Knowledge Exploration in Archaeology. Proc. AVI 2012 (pp. 681-684), ACM Press.

[5] Ricci F., Rokach, L., Shapira, B. Introduction to Recommender Systems Handbook, Recommender Systems Handbook, Springer, 2011, pp. 1-35

[6] Bilandzic, M., Venable, J. (2011) Towards participatory action design research: adapting action re-search and design science research methods for urban informatics. Journal of Community In-formatics 7(3).

[7] Orngreen, R., Pejtersen, A.M., Clemmensen, T. (2008) Themes in Human Work Interaction Design. In: IFIP International Federation for Information Processing, Volume 272, pp. 33-46.

Into The Woods

Judy Bowen, Annika Hinze, Sally Jo Cunningham
The University of Waikato
New Zealand
{jbowen,hinze,sallyjo}@waikato.ac.nz

Abstract. New Zealand forestry workers have a higher likelihood of being injured or killed in the workplace than workers in any other industry. Despite investigations, reviews and recommendations, the statistics continue to worsen year by year. We believe that in order to fully understand the reasons for this we need to undertake large-scale data collection of workers and the wider context of their working, and personal conditions. To perform such data collection in an unobtrusive manner we propose the use of lightweight, wrist-worn activity trackers. In this paper we discuss our initial plans for conducting such data-collection and the problems and challenges presented by the forestry environment.

1 Introduction

The New Zealand forestry industry has the highest fatality and injury rate of any industrial sector in NZ. Thirty-two workers have been killed since 2008 and for the past three years both death and injury rates have continued to increase. An independent review has been conducted, and results published October 2014 [1]. Data was gathered through interviews and self-reporting of all involved in the sector, from forestry owners down to the workers themselves. The report identified a number of factors such as worker fatigue, lack of training, poor health and safety cultures in the workplace etc. and made eleven recommendations. These were mostly based on the creation of processes, action groups and codes of practice to support and increase participation in training and certification for workers and contractors. However, no strategy was recommended for how the underlying causes might be identified or monitored, nor were any practical solutions for the unsafe work practices identified or questions asked as to why these continue to exist.

Existing research into reducing accidents in forestry are focussed in two key areas. The first is increasing mechanisation in the industry through improved technological advancements that will enable more machinery to be used in NZ's extreme terrain. The second is through increased health and safety and worker management. Whilst both of these have the potential to improve safety within the industry, we believe that there are additional factors that need to be studied and taken into account, we discuss this next.

2 Data Collection for Identifying Potential Hazards

A common theme that is seen in the independent review, and comes up in conversations with both workers and management in the forestry industry is that of the workers' high physical workload coupled with tiredness and fatigue. We distinguish these as tiredness being a cognitive state associated with lack of sleep and fatigue being a physical and cognitive state associated with high levels of activity. Anecdotally a forestry worker's daily physical exertion is equivalent to that of running a half-marathon. Forestry workers are working in outdoor environments where they may face extreme temperatures and weather conditions during long working hours without the necessary hydration and/or nutrition.

In order to understand the genuine effects of tiredness and fatigue in a physically demanding work environment and their contribution to accidents we believe that the first step is a large-scale data gathering exercise which enables the existence and effects of tiredness and fatigue to firstly be validated so that, secondly, we can begin to understand how we might use this knowledge to make the workplace safer. We also believe that factors outside of the work environment may be equally influential, so if for example, workers have poor sleep patterns and are already tired at the start of the working day, this may exacerbate the work-generated fatigue and tiredness.

3 Problems with Data Collection In Forestry

Trying to measure worker activity levels by observing them during the working day is inherently problematic. Known factors such as the Hawthorne Effect [2] can alter behaviours of workers who are being observed. Additionally, in high-risk work environments such as forestry, the presence of researchers can contribute to safety problems. In [3] Parker attached video cameras to forestry workers in order to try and 'observe' different work practices of novice vs. experienced workers. Initially he tried to perform the study by personally observing workers in their work environment. However, it was soon clear that what he spent most time observing was the workers trying to make sure he did not get injured or in the way – the focus of their attention switched to keeping him safe. As we are also interested in collecting data outside of working hours, for example sleep data, it is neither practical nor ethical to propose doing this via researcher observation.

In order to resolve these practical issues we investigated the use of lightweight, low-cost activity trackers (e.g. FitBit, www.fitbit.com, JawBone www.jawbone.com etc.) to collect data on the activity levels of workers, both

in and out of work, as well as sleep data. Activity trackers would enable us to collect data from large numbers of workers who may be in remote locations, in a cost-effective, efficient and accurate manner.

Our initial period of research was spent analysing various activity trackers in order to compare differences in their functionalities and accuracy in measuring different tasks (including sleeping, step counting and the effect of activities such as driving or operating a chain-saw). This analysis took place in-house and was performed by the researchers in order to make some initial decisions about which devices we should use for our first studies with actual forestry workers. Deciding which devices to evaluate is not an easy task in itself. New devices come onto the market monthly, with increasing functionalities and claims of effectiveness. At the time our first studies were performed we were analysing what were considered 'market leaders' for devices in the less than NZ$200 range (which was our maximum price point). These were devices that measured both sleep and step counting, had some analysis (based on user metrics and steps) for calorie burn, and in the case of one device (the Withings Pulse) the ability to measure heart rate.

Since these initial studies all of the leading devices now come with heart-rate measurement and more accurate calorie burn measurements, and no doubt by the time we have completed our next study these will also be out of date – in fact one of our recent experiments had been to place a heart-rate enabled tracker inside a baseball cap to see if it could still effectively measure steps and heart-rate (with the aim of looking at incorporating devices into work safety gear) – and there are now at least two commercial versions of this available on the market for less than US$200 (SmartCap from spreewearables.com and SmartHat from life-beam.com).

In order to consider both activity levels and sleep quality for workers we wanted them to wear the activity trackers 24 hours a day, but we found that many devices recorded activities such as driving – particularly on rough, unsealed roads - as steps. There is no way to identify from the data provided how much of the activity recorded is from walking and how much from other activities such as driving. We found that for most devices, operating a chainsaw to fell trees did not trigger step-counting (although the action of slicing a felled tree did).

Generally the step-counting of all of the devices was at least consistent (if not 100% accurate) with the above exceptions. Sleep-tracking was much more inconsistent and it was not clear that any of the devices or phone apps we experimented with were providing accurate data.

Once we had narrowed down the devices that seemed most appropriate, we planned a small-scale pilot study with a couple of forestry workers to see how

well the trackers would perform in our real-world scenario. There are, of course, serious ethical considerations for this sort of monitoring which includes data from outside of the work environment. Whilst it is our intention to use the data from these studies in an anonymous fashion to set safe benchmarks for sleep and workloads, increasingly these types of trackers are being used to monitor workers for performance and management reasons [4]. In order to ensure our studies are conducted in suitably ethical ways we must ensure that all parties (both workers and the management layers above them) are clear about the nature of the data we can report back and the format we will provide it in. Similarly we need to find ways of reassuring the workers themselves that this type of data will not in any way be used 'against them' to try and show they are behaving in an inappropriate way (e.g. slacking off at work or partying all weekend).

Gaining permission to conduct our studies and finding willing participants also required navigating the overly complex hierarchy of forestry owners, managers, contractors and health and safety bodies. Most workers are employed by small contractors who provide work for companies that represent forestry owners, so there is no clear-cut management layer and permission and consent is required from all of the various bodies concerned.

One final consideration that cannot be overlooked is the demographic nature of the forestry workforce (predominantly male; young; from Māori or Pacific Island heritage) compared to that of the research team (female; Welsh, German and American and over the age of 40).

4 Technical Challenges

Our initial pilot study with forestry workers consisted of two workers wearing different trackers (one Fitbit Flex and one Jawbone UP) for a period of two weeks. We visited the workers at their employer's premises to set up the study and to provide them with information about how/when to charge the devices and to agree a time when we could contact them by telephone every few days in order to resolve any problems. Their work location (a three and a half hour round trip for us) meant that we could only meet with them in person once a week. At the end of the first week we met with the workers to collect the initial data, we had been unable to contact them on the provided phone numbers during the previous week. One of the devices had a flat battery and the other was very low (neither had been charged in accordance with the provided schedule). We swapped out both of the devices for fully charged equivalents and re-iterated the importance of charging every 4-5 days. During the second week we were still unable to speak to the workers on the phone, but

found that if we texted them instead we did get a response. The data from the first week was very incomplete. One of the workers had failed to put the device into sleep mode on any of the nights and was confused by the limited LED display and the other worker had (inadvertently) put his device in and out of sleep mode several times a day and during the night; as such, it was hard to accurately calculate sleep time for either worker.

The full set of problems encountered with the three week pilot were: trackers not put into sleep mode so sleep data patchy or missing entirely; trackers not charged regularly so missing data for one or two days at a time; one tracker lost temporarily in the first week, another lost in week three for one; the charging and data connector for one device was lost and never recovered.

It is possible that problems encountered with this initial study - lost devices, inability to change modes, lack of charging – may have underlying causes other than technical challenges. We need to be mindful of the fact that whilst their manager was keen to be involved in the study and had recommended these two workers as suitable participants (and they seemed keen to participate), the reality of being monitored in this way, particularly outside of work hours, may have been uncomfortable for them. We are not suggesting that there was any attempt to 'sabotage' this study, but it is certainly a possibility that may occur, particularly in larger studies. The issue of data anonymity was pertinent to this first study as we were only using two workers as our sample study. While their manager was keen to see the data to understand the workloads they were under, it was difficult for us to provide anything more general than the average step count per day for the workers combined without breaking privacy constraints.

5 Our Next Investigations

Despite the problems we encountered with our pilot study it was informative in enabling us to set and manage expectations for future studies. One of the things that has become clear is that the practice of monitoring worker's sleep patterns (even for short periods of time as an investigative option) is fundamentally problematic both from a practical as well as an ethical perspective. In order to be able to use such data as an indicator of potential danger in the working environment it would require both a long-term initial data gathering period (to set base-lines for individual workers) as well as on-going monitoring so that changes from the baseline could be identified. Again, the ethics of this sort of monitoring (even if the intention is to improve safety) is questionable.

Of most interest to us is the effect that fatigue and tiredness have on workers, particularly their response times, as this is what may enable detection of potentially hazardous situations. Our next studies, therefore, are focussed on looking at ways of measuring these response times and understanding the effect that fatigue may have. For example if a worker is performing a particular task which leads to fatigue, which in turn affects their response times, this might be used to better inform forestry management as to how tasks should be allocated throughout the day.

We are currently undertaking a survey of around 30 forestry workers to gain some information about how they feel they are affected by fatigue and tiredness. This will be followed by another small-scale pilot study where we will use a response-measuring app to record worker response times at three different points during the day, whilst also recording activity (steps) and heart-rates during the working period. The researcher for this project (a male, MSc student who lives within a half hour drive of the workers' location) will meet with the workers several times a day (for a period of one week) to perform the response tracking and collect the data from the activity trackers. The researcher will also be responsible for charging the trackers at night when necessary to ensure that data can be collected each day. This will help mitigate some of the problems we saw with our previous pilot study.

We have, however, already run into our first problem with this study, that of ensuring the researcher has correct personal liability insurance to be covered whilst in the forestry environment. No doubt as the study progresses we will uncover further issues of interest. In addition, one of the Health and Safety organisations involved in the forestry domain has offered to supply an additional quantity of activity trackers to be used as part of this study so that a larger number of workers can be surveyed. This, however, raises questions about data privacy (who do the devices 'belong to' and who has access – and ownership – of the data captured) and is yet another hurdle we need to navigate before this study begins.

6 Conclusions

In this paper we have briefly described an on-going research project which aims to perform large-scale data collection of forestry workers. We have also outlined the main problems encountered with this sort of research, which includes technical, ethical and sociological factors. We have identified some 'solutions' which we have adopted for our next proposed studies (for example the use of an in-situ researcher on a daily basis to help mitigate some of the

technical challenges), however we also expect to uncover further challenges as we progress with this, and subsequent studies.

It is likely that there are other research environments which contain similar challenges and we look forward to the opportunities of discussing these and sharing solutions.

References

[1] Adams, G., Armstrong, H., Cosman, M., Independent Forestry Safety Review – An agenda for change in the forestry industry. October 2014. Available from: http://fica.org.nz/wp-content/uploads/2014/10/IFSR-ReportSummary-Web.pdf (May 2015)

[2] McCarney R., Warner J., Iliffe S., van Haselen R., Griffin M. and Fisher P. The Hawthorne Effect: a randomised, controlled trial. British Medical Council Journal of Medical Research Methodology 7:30. doi:10.1186/1471-2288-30. PMC 1936999.PMID 17608932.

[3] Parker, R., Technological advances in the analysis of work in dangerous environments: Tree felling and rural fire fighting. PhD thesis, Massey University, Palmerston North, New Zealand (2010)

[4] Sagan, A., Employee tracking apps raise worker privacy questions, CBC news article, May 19th, 2015, retrieved from: http://www.cbc.ca/m/news/technology/employee-tracking-apps-raise-worker-privacy-questions-1.3075673 (May, 2015)

Designing for the Factory: UX Prototyping for the Cleanroom

Roland Buchner, Verena Fuchsberger, Astrid Weiss, Manfred Tscheligi
Center for Human-Computer Interaction
University of Salzburg
Sigmund-Haffner Gasse 18, Austria
Firstname.lastname@sbg.ac.at

Abstract. In our work on facilitating human-computer interactions in the (smart) factory through contextual interfaces, we addressed the very specific user group of maintainers in a factory. Maintenance activities need to be flexible, quick, and cooperative, which, in turn, leads to specific challenges in research and designing interfaces for them. We present a case study that is based on a two-step user-experience prototyping process of a contextual inquiry and a participatory design workshop with actual maintainers and their co-workers (to include cooperative aspects). We identified the need to consolidate information originating from various interfaces and ubiquitous systems deployed throughout the factory in order to improve their workflow. On the other hand, design needs to facilitate awareness for broken and serviced equipment. To address these needs we developed a User Experience (UX) prototype that consists of a mobile interface for maintainers to enable tailored information presentation and situated displays intended to raise equipment awareness of broken equipment for co-workers in the cleanroom. In this position paper we discuss our approach and highlight the need to further research mobile interactions and situated displays in such a dynamic context.

1 Introduction

In the context of a semiconductor factory low production costs and reduced defect rates are essential for the success of a factory. The goal of such companies is to strive for a "zero-defect" production, as many of their products are used in safety critical contexts (e.g., automotive, aerospace). To reach this goal a rising number of factories these days is automated, however, humans still a play a crucial role in the production. Maintainers are among these important people as they are responsible to service the equipment in the cleanroom to guarantee a high quality production. To reach such high quality productions, research has already shown that by means of mobile interfaces (e.g., [3]) or situated interfaces (e.g., [6]) production processes in the factory could be improved. Regarding maintenance activities, the use of augmented interfaces in the factory context has shown promise and was researched thoroughly lately (e.g., [7]). However, most of the research only considered the single user in the factory, collaborative aspects have mainly been

overlooked. In our research of the semiconductor factory we found that cooperative aspects need to be considered as they affect the performance of how equipment is being serviced [4], e.g., to raise the sift leads' awareness of defective equipment. Research (e.g., [2], [8]) has shown that by providing the proper communication channels cooperation between maintainers in the factory could be improved. However, with our approach we do not only concentrate on a single user group (e.g., maintainers) we also want to keep fellow workers who are not maintainers (e.g., shift leads, operators) up-to-date of the equipment activities.

Furthermore, we identified pitfalls of interfaces in the factory, e.g., a lack of tailored information and limited acceptance of the deployed interface [4], which we also want to address in this paper.

Our approach is to develop experience prototypes that help to get a better understanding of the interaction experience as the cleanroom is a very difficult and complex context to research. They allow all stakeholders (e.g., maintainers, shift leads, operators) to gain first-hand experiences of the prototype's existing and future conditions by actively engaging with it [1]. In so doing, we present a case study where we describe a User Experience (UX) prototyping process that consists of a Contextual Inquiry (CI) [4] to develop an initial UX concept and a follow-up workshop with maintainers and co-workers to iterate and to discuss collaborative aspects (e.g., knowledge transfer) of the prototype. Based on the results of these two steps a UX prototype was developed that consists of a mobile interface in conjunction with situated displays. This combination supports different degrees of mobility and provides a smooth change of activities - from individual to collaborative ones, and from shared to public information and vice versa [5]. We aim at supporting personalized information presentation and information consolidation to reduce the information overload. By means of situated displays we want to visualize department specific information to reach a broader group of people to increase equipment awareness and to support cooperation between maintainers and co-workers (e.g., shift leads).

2 Contextual Analysis and Participatory Workshop

The **first** step was to gather the requirements and to understand the context, maintenance work routines, and maintenance tasks. Therefore, we (two researchers) conducted a Contextual Inquiry (CI) in the cleanroom of a semiconductor factory for three days and a two hours follow-up workshop with maintainers the fourth day [4]. The researchers gathered interview data using a dictaphone and handwritten notes. The interview data was transcribed

and qualitatively analyzed using NVivo[1] to better understand maintenance working procedures and processes. In the following we present findings of the CI that we categorized in three topics, namely: *interface usage, active information seeking and notification, and equipment handling*. These three topics were used as design rationale to conceptualize a first UX prototype.

Regarding *interface usage*, we identified about 15 different interfaces that are used by maintainers to complete their work. These interfaces often share the same or similar data, which create an information overload. The factory also has deployed terminals that run a central coordination interface throughout the factory that is considered to be a unified interface access to all different kinds of equipment items. Maintainers use it to check if other equipment items of interest are offline or to document their work. Their interface usage differs greatly and depends significantly on their working experience in the factory, and is furthermore historically influenced by their pre-experiences when new interfaces were deployed in the department.

Regarding *active information seeking* maintainers often look at terminals whenever they pass one to be up-to date on the equipment status. We also observed that maintainer got *notified* about the equipment by co-workers (e.g., shift leads, operators) or by the automated process control system.

Regarding the *equipment handling*, we found that the equipment is prioritized and assigned to specific maintainers at each shift change by the shift lead in their department. The rationale is that there are lot of different equipment types in the factory that are needed for the wafer production. Since the equipment is very complex maintainers are often specialized to only a specific type of equipment. If such prioritized equipment breaks then the responsible maintainer needs to check on it. During a shift cycle the priority of the equipment might change and therefore maintainers and co-workers need to be aware of such changes.

To develop a system that supports maintenance, the information needs to be tailored for maintainers and co-workers to reduce the amount of information. To improve the knowledge transfer and equipment coordination the equipment priority and maintenance activities need to be transparent and accessible to a greater group of people in the cleanroom. Based on the CI an initial concept for a mobile interface was developed, which displays personalized information, and a situated display to present equipment information department wise.

---

1 NVivo is a software for qualitative data analysis

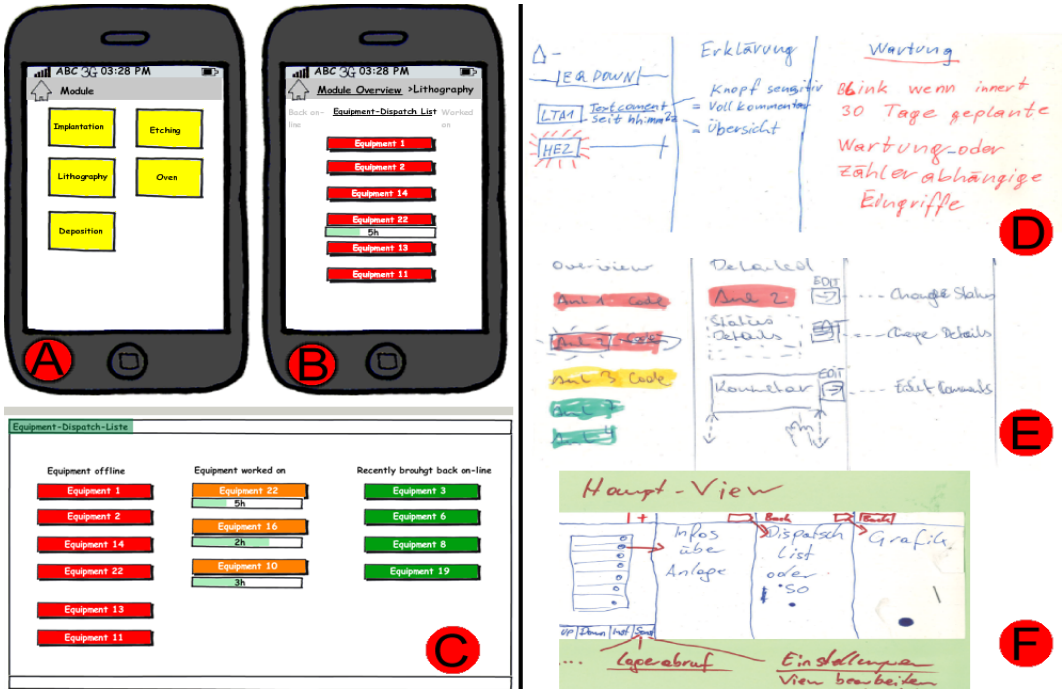


Figure 1. The letters A, B, C show the initial concept based on the results of the CI that was presented in the workshop. The letters D, E, F show images from the prototypes that were drawn in the workshop by the participants.

In the **second** step we conducted a participatory design workshop were we presented and discussed the initial UX concept (see Figure 1 A-C). The ideas consisted of a graphical and a scenario-based description of the mobile interface and the situated display - the former showing how personalized information could be displayed and the latter showing equipment related information to foster cooperation. The goal of the workshop was to discuss the relevance of information that needs to be displayed on the mobile interface and on the situated display. After the presentation the participants were asked to write down key words concerning what they thought was important for a maintenance tool and briefly explained their ideas. Afterwards, the key words were grouped and attached onto a wall. For example, in the discussion of the initial concept with the participants it became clear that a progress bar was not wanted (see Figure 1 B and C). It was stated by the participants that time predictions for maintainers were too difficult, even for the very experienced ones, and maintainers would feel being monitored during work. A missing information in the initial concept, which the participants claimed as important, was the information of who did changes to the equipment (e.g., who changed the priority of an equipment item), which was added later in the finial concept. After the presentation and the discussion the participants were asked to work in groups of two to develop paper prototypes based on their knowledge and experience in the maintenance area. To give them an idea of what prototyping

is, we briefly showed them what paper prototypes could look like, and we encouraged the participants to create their own paper prototypes (see Figure 1 D-F). Based on the outcome of the workshop the initial concept was iterated and an interactive User Experience (UX) prototype was developed, which is presented in detail in the next section.

3 The User Experience Prototype

The context of the semiconductor factory poses several challenges for contextual research. The cleanroom is an almost dust-free area with a 24/7/365 production cycle where the workers have to wear full body suits to reduce particle pollution. These challenges make it difficult to develop and to evaluate novel systems for this context, as the production must not be affected. That is why we chose to develop experience prototypes as they help "to understand, explore or communicate what it might be like to engage with the space or system" [1].

Our UX prototype consists of a mobile interface in conjunction with situated displays that are intended to support the maintenance workflow and to improve the transparency between different actors in the factory. The prototype addresses three aspects that we consider as relevant: reduction of the information, presentation of personalized equipment information, as well as fostering cooperation and support equipment coordination.

An overview of the prototype is given in Figure 2, which shows the screens of the mobile interface (A-D) and the situated display (E). To address the issue of information overload, different equipment states (offline, at work, on hold) are presented on the mobile interface as tabs as a first layer of information reduction.

Figure 2 A shows the default screen of all the equipment that is of importance for the maintainer. This is a personalized list, where the maintainer can add or remove equipment items. The screens B to D show only a subset of information that is available on screen A. Screen B shows equipment that needs to be serviced (defective equipment or where the service on the equipment is due) colored in red. For defective equipment the maintainer needs to know who is responsible and who set the priority. In screen C the equipment that is being worked on (equipment that is cleaned, serviced or repaired) is colored in yellow. Here, the maintainer knows who is working on the equipment and since when.

Screen D shows the equipment that is set on hold colored in grey (e.g., a broken part needs to be ordered from another company). Equipment that is colored in green is in a working state (actually processing or waiting for

wafers to process) and is only visible in screen A. To reach a high productivity it is important that the time the equipment is offline is as short as possible. However, depending on how many equipment items are offline, there might not be enough maintainers to work on all offline items, therefore the equipment is being prioritized, thus, maintainers need to know the equipment priority. To visualize the priority each tab lists the equipment from top (most important) to bottom (least important). The information on the equipment that needs to be visible at a glance is: the exact point in time of the equipment state change (e.g., online to offline), upcoming or current events, who set the priority or changed the equipment state, who is the responsible maintainer, and who is working on the equipment. The mobile interface provides maintainers with up-to date information of "their" equipment. They do not need to check on terminals all the time when they pass one to be up-to date on the equipment.

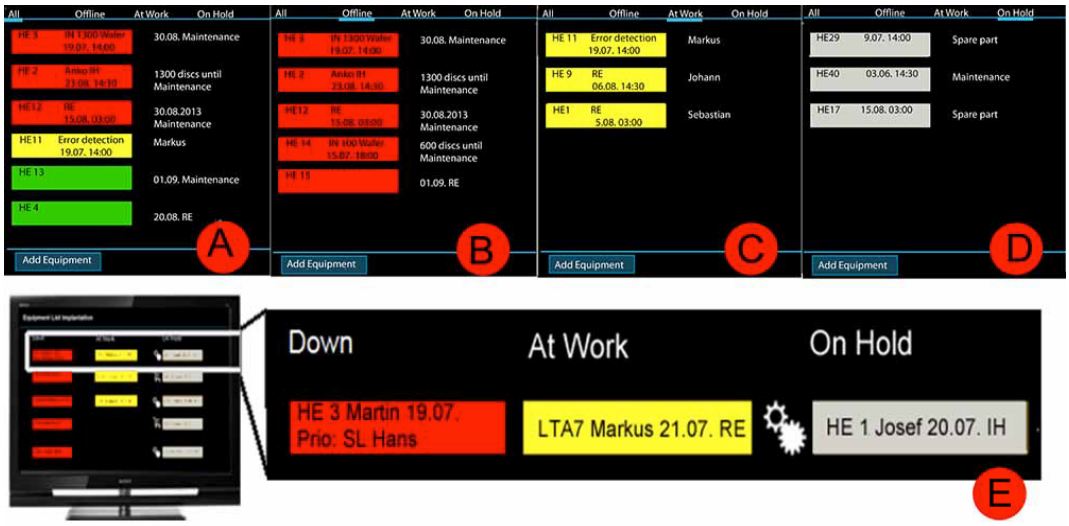


Figure 2. The mobile interface (A-D) and the situated display (E) after the CI and the workshop

The situated display was developed (see Figure 2 E) with the purpose to present information on defective equipment, equipment that is being worked on, and equipment that is on hold, in an intuitive and comprehensible way for shift leads, operators, and fellow maintainers. In contrast to the mobile interface the situated display does not show personalized information – the information is department specific and not responsive to user input, but with the same degree of detail. An additional information that is not shown on the mobile interface is the icon next to the on-hold state, which represents what needs to be done here (e.g., gearwheels indicate a cleaning or maintenance task). Due to the abstracted information visualisation of the situated display in comparison to the central coordination interface workers only get the

information of the equipment that needs to be serviced or is being serviced. As all workers in the cleanroom need to have a high productivity it is in their interest that the equipment is online and ready to process, therefore it is important to know its actual state at a glance.

4 Conclusion

In this paper we presented a case study of how findings from a user experience (UX) prototyping process in the challenging context of a semiconductor factory informed the development of a UX prototype (a combination of mobile interfaces with situated displays) to support maintenance activities and cooperation between co-workers in the cleanroom. The access to the cleanroom is very restricted and regulated with strict rules (e.g., no running, mobile phones need to be put in plastic bags) as the production of wafers is very sensitive and must not be disturbed by anyone. However, by means of a contextual inquiry (CI) [4] to gather requirements and the participatory workshop with actual maintainers we were able to carry out a UX prototyping process in such a challenging context in a feasible and fruitful way. The CI proved to be a reliable method to get an understanding of maintenance work in the factory, building the basis for the initial UX prototype concept. The workshop also showed its value as different participants got to know different opinions and interests when it comes to maintenance. Based on the workshop results the UX concept was iterated. However, to study user experience and cooperative aspects of the prototype, we need to explore a prototypical implementation of the system in the semiconductor factory in comparison with actual factory on-site solutions in order to continuously advance human-computer interactions in smart factories in various ways.

Acknowledgements

The Financial support by the Federal Ministry of Economy, Family and Youth and the National Foundation for Research, Technology and Development is gratefully acknowledged (Christian Doppler Laboratory for “Contextual Interfaces”).

References

[1] Buchenau M., Suri JF. Experience prototyping. DIS ’00, ACM, NY, USA, 424–433,

[2] Fallman, D. Wear, point, and tilt: designing support for mobile service and maintenance in industrial settings. In: DIS' 02
[3] Gehring, S., Löchtefeld, M., Schöning, J., Gorecky, D., Stephan, P., Krüger, A., Rohs, M. Mobile product customization. In: CHI EA '10
[4] Kluckner, P.M., Buchner, R., Weiss, A., Tscheligi, M. Collaborative reporting tools: An analysis of maintenance activities in a semiconductor factory. In: CTS 2013 . pp. 508–515
[5] Luff P. and Heath C. Mobility in collaboration. In *Proc. of* CSCW '98. ACM, New York, USA, 305-314
[6] Meschtscherjakov A, Kluckner P, Pöhr F, Reitberger W, Weiss A, Tscheligi M, Hohenwarter K, Oswald P. Ambient persuasion in the factory: The case of the operator guide. In: Advanced Semiconductor Manufacturing. 2011.
[7] Vartiainen, E., Boberg, P., Qvarnstrom, O., Brönmark, J. Applying augmented reality to industrial settings. In: Proc of the 4th Augmented Human Int. Conf. pp. 244–244. New York, NY, USA (2013)
[8] Väätajä, H., Janhunen, J. Supporting mobile maintenance work - needs and requirements. In: Industrial session at NordiCHI2012, Copenhagen (2012)

Experience Design and Positive Design as an alternative to classical human factors approaches

Michael Burmester[1], Katharina Maria Zeiner[1], Magdalena Laib[1], Cristina Hermosa Perrino[1], Marie-Luise Queßeleit[2]
[1]Stuttgart Media University, Nobelstr. 10, D-70569 Stuttgart
[2]SIC! Software GmbH, Ferdinand-Braun-Straße 1, D-74074 Heilbronn
[1]burmester|zeiner|magdalena.laib|ch119@hdm-stuttgart.de
[2]marie-luise.quesseleit@sic-software.com

Abstract. The classical human factors approach of designing for demanding work environments is to design for optimal task support, reduction of stressful conditions, and eliminating usability problems. Following experience design and positive design we investigated the feasibility of increased positive experiences for users in such environments. Two interaction concepts for field sales personnel were developed to show that it is possible design for positive social relationships to customers and for gratitude for contributions in customer relationship management systems. Possible positive consequences of these interaction design concepts are discussed.

1 Introduction

This study is about field sales personnel selling tools and materials to craftsmen by visiting them. This is a very demanding work environment because: (1) In most cases sales personnel are non-experts in the field of knowledge of their customers. The craftsmen are always more competent, because sales people tend to supply different craft businesses (e.g. painters, drywall installers, plumbers). (2) Their customers often feel disturbed in their work and are annoyed when they are being visited. (3) Many sales companies put a lot of pressure on the sales personnel by setting strict, often hard to reach sales targets. This is worsened by the increasing number of alternatives open to customers (especially online and hardware stores) and the fact that most customers buy products from several sales companies. (4) Sales personnel have many appointments per day and have to drive over long distances by car. (5) An important success factor for selling products is a positive social relationship to customers. This is in contradiction with the negative attitude of many customers and the stress inducing work conditions. Sales work is emotion work [1]. (6) Sales personnel are using product catalogues with 200.000 to 400.000 products. Some are still using printed catalogues and some

digital catalogues on tablet PCs. Digital catalogues are often nothing more than PDF files or simple applications providing search functionality. Both catalogue solutions are difficult to handle and do not support presenting of related products or special offers. The conclusion is, that the environments field sales work in are mentally demanding because of the juxtaposition of stressful work conditions (time and success pressure, missing knowledge, insufficient tools, and difficult customers) and the task to set up positive and lasting relationship with their customers. An indicator of the demanding work environment is the high fluctuation of employees.

The question is, whether it is possible to design applications to support the field sales force. The traditional approach of design for good usability would be to analyse the context of use, including user characteristics, task structure and environmental requirements [2]. Of course, this is a basic way of designing applications for work contexts and ensures that technology is designed according to the requirements of tasks and work environment. As Desmet and Hassenzahl propose, the inherent notion is to design technology as a problem solver and call it "problem-driven design" [3]. As is, this approach is not amiss, but the focus is on avoiding or neutralising the negative and misses possibilities to design for positive experiences ("possibility-driven design", [3]). Taking the idea of good design even further, research on user experience focuses on positive emotional experiences when interacting with software. Therefore, our research question was, whether it is possible to design for optimal task support with a clear focus on the design for positive emotional experiences in this demanding and stressful work context.

2 Theoretical aspects and related work

Steger and Dik [4, p. 131] point out that work affects how we understand our lives and our environment. Technology is an integral part of most modern jobs, therefore designing for positive user experience of that technology might contribute to more fulfilling work environments. Outside of work environments, the experience of technology and its contribution to well-being (e.g. Positive Design [5]) has been explored repeatedly,

Hassenzahl [6, p. 12] defines User Experience (UX) as "as a momentary, primarily evaluative feeling (good-bad) while interacting with a product or service. Good UX is the consequence of fulfilling the human needs for autonomy, competency, stimulation (self- oriented), relatedness, and popularity (others-oriented) through interacting with the product or service (i.e., hedonic quality)." A positive experience through and with the use of technology thus emerges from designing according to psychological needs. A

second approach is followed by Desmet and Pohlmeyer with their framework of Positive Design [7]. It has the potential to contribute to subjective well-being by addressing three different domains: Pleasure, personal significance and virtue or all three of them. Positive Design is based on the theory of Positive Psychology [8] which investigates what aspects contribute to a positive emotional state and defines strategies to achieve it ([9], [10]).

Similarly, positive UX may contribute to positive changes in work contexts [11], [12]. Kohler and colleagues, for example, demonstrated that giving a simple feedback about achieving subgoals of a task increased UX and heightens motivation [13].

3 Concepts for software design for positive user experience

3.1 Research question and methodology

Here we asked, whether it is possible to design a catalogue system which facilitates positive emotional experiences by improving search for products and presentation of related products. The lack of domain knowledge of the sales staff should be compensated. Sales staff should experience competency during the sales pitch. In addition possibilities for positive UX should be implemented in a customer relationship management system (CRM) for sales staff in the field.

We applied a design process for positive UX [14]. It consists of the following activities: analysing the work context, collecting existing positive experiences using experience interviews, designing user experience concepts based on interview insights and design strategies of experience design [15] and positive design [7], followed by prototyping of the experience concepts, as well as, evaluating them with a focus on analysing positive and negative experiences [14], [16].

3.2 Experience interviews

Experience interviews are in-depth interviews probing for positive experiences in the work contexts. Semi-structured interviews were conducted face to face or via telephone: Participants were asked to describe their daily routine, such as customer meetings, and the usage of artefacts like catalogues and software. After that the social relationship between the participant and others were discussed. Finally participants were asked to report positive experiences in the context of work, in other words situations when they felt happy and fulfilled.

Experiences included those independent of as well as linked to technology were explored deeper, using the laddering-technique [17].

Seven male sales representatives, working in the following different companies and sectors, participated in the interviews. The interviews showed, that sales related jobs can differ vastly. We specifically chose both positive and negative work contexts to gain a better understanding of the heterogeneity of the field. The concepts we developed based on these interviews were, however, specifically aimed at those working in the emotionally more demanding settings. A short description of the different organisational cultures derived from the interviews is given:

- A sports car manufacturer: The corporate culture was described as cooperative. Every salesman is assigned to a territory and works on commission basis. Sales representatives are encouraged to build trust-based and long-term relationships. Closer relationships, however, are regulated through the company's code of conduct.
- A humidifier manufacturer and industrial humidification specialist: The corporate culture was described as professional. Every salesman is assigned to a territory and while they earn a base salary, commissions are still an important factor. Relationships with customers were characterized as business related.
- A regional heating construction company with a network of 1100 companies: The corporate culture was described as interpersonal. Salesmen earn a base salary and if the targeted sales figures were achieved, they receive a share of the turnover. Salesmen are encouraged to strive for positive and friendly relationships.
- A regional trading company in the fields of tools, hardware, assembly technology, and workwear: The corporate culture was described as cooperative. Salesmen earn a fixed salary. The interviewees reported building positive and friendly relationships with customers.
- A large trading group focusing in fixing and assembly technology: The corporate culture was described as competitive, with the company striving to enhance customer relationships. Therefore, each sales representative tends to work with the same customers. While salesmen earn a base salary, their pay is also linked to commissions.

The gathered anecdotes were analysed to discover factors, which affect their lives positively, while finding common structures of positive experiences.

After a qualitative content analysis, we elicited the following positive aspects and experiences: Sales representatives enjoy the wide independence and freedom when planning their own daily routine (3 participants (p)). It is

exciting for them to deal with the rich variety of different tasks, situations, and customers of different industrial sectors (3 p). A good customer relationship is beneficial for achieving success, but it is also pleasant in itself. Sales representatives describe themselves as extrovert and communicative, and they report that they get positively excited by the contact to their customers. A very important type of experience is that they enjoy supporting their customers in their projects (6 p). Several situations were mentioned as leading to a feeling of success: the sales representative provided a solution or solved a problem (6 p), closed a sale (6 p), increased the sales (5 p), satisfied the customer and earned gratitude for his effort (5 p). The personal relationship between the sales representative and the customer was also mentioned as pleasant - they talk often about private issues or hobbies and in doing so they can establish a good relationship over time (7 p).

3.3 Interaction Design Concepts for supporting positive experiences

One of the tools used by the sales personnel is a customer relationship management system (CRM). The idea is, that the information concerning customers is stored there and available to the entire sales team. The CRM passes helpful information (business and personal information) to colleagues. Especially personal information supports communication, introduction, and small talk. Receiving helpful information might lead to a feeling of thankfulness towards the person entering this information. Inspired by the results of positive psychology that gratitude is positive for the person expressing gratitude and for the person receiving it, we designed a possibility for expressing thankfulness for helpful information. A special “that was helpful” button associated to an information input field can be used to express thankfulness. An important aspect is, that the interaction was visible only to the two users involved in order to prevent a competitive situation. Following a positive psychology approach this design has the potential to support subjective well-being and social relationships [18], [19].

The second concept is based on the experiential interview insight, that working together with the customers on their project is a very positive and satisfactory experience for the sales personnel. The concept project collage (fig. 1) offers the possibility to plan a project in co-operation between the sales person and the customer. This is done by identifying project activities and then collecting the relevant materials and tools. In fact, entering project activities starts searching the catalogue database and presents materials and tools related to project activities. The sales representative is put into the role of a moderator

for the project planning process and the customers is assigned the role of the expert for their projects and work. Furthermore, the project collage stores projects as templates, so that it is easy to set up comparable projects and refine them. This design should increase the possibility to experience competency because they are now part of the planning process with their customers. This addresses the need for a feeling of competency [20] in the interaction with customers.

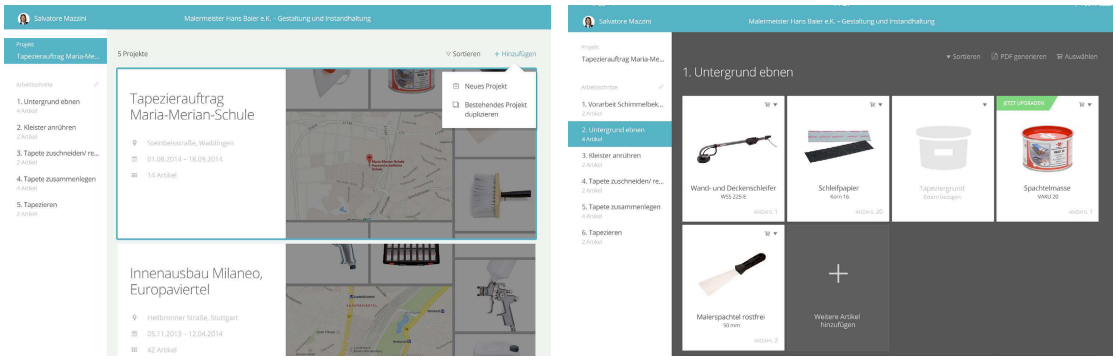


Figure 1: UX concept project collage (left: list of project templates; right: products presented to the second project activity ‘smoothing a wall’)

4 Discussion

The two concepts demonstrate that it is possible to design for positive emotional experiences in a very demanding work environment supporting the fulfilment of psychological needs and using strategies for well-being.

These concepts are currently under empirical evaluation. A problem here is the problematic access to this special user group causing research delays.

Research of positive psychology at work showed that supporting social interactions with customers [20] and gratitude [18], [19] are a source for positive experiences. According to the broaden and build theory [21], positive emotional experiences have several positive effects like increased engagement and optimism, developing of competences and social relationships, more flexible thinking, acting and problem solving, as well as reducing the effects of negative experiences and increasing resilience.

The next step in our research will be evaluating whether the concepts generate positive experience during interaction. After that the positive consequences for work in demanding work environment and the interaction with digital tools can be investigated.

Acknowledgements

Special thanks to the students working on the experience concepts: Clemens Bergmann, Kathrin Buess, Carina Gaßner, Kristin Haasler, Franzika Joos, Anja Koppenhöfer, Domink Pfeifer, Steffen Schäfer, and Marc Walter.

References

[1] D. Zapf, "Emotion work and psychological well-being A review of the literature and some conceptual considerations," *Hum. Resour. Manag. Rev.*, vol. 12, pp. 237–268, 2002.

[2] ISO 9241-210, "Ergonomics of human-system interaction - Part 210: Humanacentred design for interactive systems," 2010.

[3] P. M. A. Desmet and M. Hassenzahl, "Towards happiness : Possibility-driven design," in *Human-computer interaction: The agency perspective*, M. Zacarias and J. V. de Oliveira, Eds. New York, NY: Springer, 2012, pp. 1–27.

[4] M. F. Steger and B. J. Dik, "Work as Meaning: Individual and Organizational Benefits of Engaging in Meaningful Work," in *The Oxford Handbook of Positive Psychology and Work*, P. A. Linley, S. Harrington, and N. Garcea, Eds. Oxford: Oxford University Press, 2013, pp. 131–142.

[5] P. Desmet and A. E. Pohlmeyer, "Positive Design : An Introduction to Design for Subjective Well-Being," *Int. J. Des.*, vol. 7, no. 3, pp. 5–19, 2013.

[6] M. Hassenzahl, "User experience (UX): towards an experiential perspective on product quality," in *Proceedings of the 20th International Conference of the Association Francophone d'Interaction Homme-Machine*, 2008, pp. 11–15.

[7] P. M. A. Desmet, A. E. Pohlmeyer, and J. Forlizzi, "Special Issue Editorial : Design for Subjective Well-Being," *Int. J. Des.*, vol. 7, no. 3, pp. 5–7, 2013.

[8] M. Seligman and M. Csikszentmihalyi, "Positive psychology: An introduction.," *Am. Psychol.*, vol. 55, no. 1, pp. 5–14, 2000.

[9] S. Lyubomirsky, *The how of happiness: A scientific approach to getting the life you want*. London: Penguin, 2007.

[10] M. E. P. Seligman, *Flourish: A Visionary New Understanding of Happiness and Well-being*. New York: Free Press, 2011.

[11] K. Klöckner, H. Schmitt, and M. Garst, "Schlussbericht FUN - (F) UN of (U) se für Geschäftsa (N) wendungen," 2009.

[12] S. Harbich and M. Hassenzahl, "Beyond task completion in the workplace: execute, engage, evolve, expand," in *Affect and Emotion in Human-Computer Interaction*, C. Peter and R. Beale, Eds. Berlin Heidelberg: Springer, 2008, pp. 154–162.

[13] K. Kohler, S. Niebuhr, and M. Hassenzahl, "Stay on the Ball! An Interaction Pattern Approach to the Engineering of Motivation," in *Human-Computer Interaction – INTERACT 2007*, 2007, vol. 4662, pp. 519–522.

[14] M. Burmester and R. Tille, "Travel Experience - Erlebniszentrierte Gestaltung neuer Medien für Reisende.," in *Usability Professionals 2013*, 2013, no. 2012, pp. 146–151.

[15] M. Hassenzahl, *Experience Design: Technology for All the Right Reasons*. Breinings-ville: Morgan & Claypool, 2010.

[16] M. Burmester, M. Mast, K. Jäger, and H. Homans, "Valence method for formative evaluation of user experience," in *Proceedings of the 8th ACM Conference on Designing Interactive Systems - DIS '10*, 2010, p. 364.

[17] T. J. Reynolds and J. Gutman, "Laddering theory, method, analysis, and interpretation," *J. Advert. Res.*, vol. 28, no. 1, pp. 11–31, 1988.

[18] S. B. Algoe, J. Haidt, and S. L. Gable, "Beyond reciprocity: gratitude and relationships in everyday life.," *Emotion*, vol. 8, no. 3, pp. 425–429, 2008.

[19] T. Lamas, J. J. Froh, R. A. Emmons, A. Mishra, and G. Bono, "Gratitude Interventions," in *The Wiley Blackwell Handbook of Positive Psychological Interventions*, A. C. Parks and S. M. Schueller, Eds. Chichester: Wiley & Sons, 2014, pp. 3–19.

[20] S. Kiffin-Petersen, S. A. Murphy, and G. Soutar, "The problem-solving service worker: Appraisal mechanisms and positive affective experiences during customer interactions," *Hum. Relations*, vol. 65, no. 9, pp. 1179–1206, Aug. 2012.

[21] B. L. Fredrickson, "The broaden-and-build theory of positive emotions.," *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, vol. 359, no. 1449, pp. 1367–78, Sep. 2004.

Challenges for Action Research on HWID in Activity Based Workplaces

Åsa Cajander[1], Gerolf Nauwerck[1], Thomas Lind[1], Marta Larusdottir[2]
[1]Uppsala University, BOX 337, SE-751 05 Uppsala
[2]Reykjavik University, 101 Reykjavik, Iceland
{asa.cajander;thomas.lind}@it.uu.se
gerolf.nauwerck@uadm.uu.se; marta@ru.is

Abstract. In this paper we discuss challenges when conducting action research and formulating research projects in a fairly volatile organisational setting. This is done against a background of Human Work Interaction Design (HWID) research in new office environments and Activity Based Workplaces (ABW). While the concept of ABW is not new, presently there is a growing trend towards reconstruction of office work-places where the concept, in one form or the other, is having an impact on office design–and the design and use of ICT in the workplace. This paper will present experiences from the planning phase of a project in the area of HWID and ABW, and discuss the different roles and situations that needed to be handled as an action researcher.

1 Introduction

In this paper we present experiences from the planning phase of an action research project in the area of Human Work Interaction Design (HWID) and new ways of working (NWoW), or more specifically activity based work (ABW). Studying how ICT and especially interaction design can support employees in ABW is an emerging field. While ICT and office design have a long history as there have been numerous pioneering efforts in establishing NWoW, recent developments in ICT, prominently regarding mobility and ubiquitous computing, alongside high profile new offices such as Googleplex, contribute to a global trend in restructuring office design and office work (Bodin, 2010).

An opportunity to study the phenomena of NWoW and HWID presented itself to the research group as a large Swedish government organisation (GovOrg), that the group had worked with earlier, suggested that a new office building white collar workers was to be based on the concept of ABW. An action research project was initiated, for which the planning phase received funding as a separate planning project. Although the field is highly relevant and the initial idea in many ways still stand firm, it quickly became apparent during the planning phase that the high stakes involved in the change as well

as the internal dynamics within the organisation presented serious challenges for the continuation of the project.

In this paper we will reflect upon the challenges presented to the initial project idea and to discuss possible strategies for continuing the research, as we believe this to be a relevant field for HWID research.

2 Background and Theory

2.1 Activity Based Workplaces

The physical work environment in all its variety is a factor influencing office workers (Bodin, 2010). Recently, flexible offices and more specifically management philosophies such as the idea of activity based working (ABW) has attracted the interest of organisations – notably in Australia, the Netherlands and Sweden – as well as researchers. The concept of ABW is rather complex and ambiguous, one description states that employeers "whilst in the office, can choose an activity-based workstation that best suits the activity at hand from a functional perspective and also matches with the employees' preferences." (Appel-Meulenbroek et al., 2011, p. 123). One important consequence of this is that it makes possible a significant reduction of traditional workstations and also office space. Space is here seen as a mediating factor between people and ICT. Indeed, the strategic use of corporate space is seen as the necessary, though not sufficient, factor in empowering the workforce and ameliorating many of the downsides of computer-supported work (van Kotsveld & Kamperman, 2011). A part of the concept is the increased use of ICT in support of both employee mobility and facility monitoring (Ianeva et al., 2015). Not surprisingly, the ICT industry itself is one of the major proponents of this new way of working (Gates, 2005). Yet, while ABW is proposed as a solution to the problems associated with open plan offices, e.g. high cognitive stress, research on ABW and cognitive stress is inconclusive (Too & Harvey, 2012, De Been & Beijer, 2014). In a seminal paper, Humphry (2014) traces the origins of this seemingly new "anywhere, anyplace" work rhetoric and expose some of its inherent paradoxes, notably e.g. how these images of newness contribute to the conservation of old work patterns.

As ICT in the workplace thus become more and more embedded and pervasive the scope for HCI and HWID broadens. The pioneering works of Alexander (2006) have argued for the application of usability concepts in the field of facility management, however the field seems to have attracted limited attention so far (Rasila et al. 2010). Instead, the most common tool seems to

be variations on the model developed by Laing et al. (1998), where the amount of face-to-face interaction is contrasted with the amount of job autonomy (resulting in a matrix of four basic office types: the hive, the cell, the den and the club). While this broad categorisation of work can aid planning, other techniques, such as personas, could provide a deeper understanding on how to improve the quality of work and the work environment.

2.2 Action Research

Action research has been our research approach for many years at the HCI group at Uppsala University. In action research projects, the role of the researchers is twofold: they are studying a particular phenomenon in practical settings (the research part) and they are actively improving the situation together with the participants (the action part) (McKay and Marshall, 2001). In many of our studies, we have focused on work environment and how ICT can support white collar workers in their daily activities (e.g. Sandblad et al, 2003; Janols et al, 2014, Cajander et al, 2014). Action research is a suitable research methodology for organisational change such as the introduction of new work patterns and research as it has the dual aim of solving research questions at the same time as solving problems in practice (McKay & Marshall, 2001) and to make a change (Oates, 2006). The idea is that the researchers, together with the practitioners, combine their different experiences and knowledge in order to solve a particular problem, as well as developing theory around this problem. Rapoport (1970) defines action research as an approach that "aims to contribute both to the practical concerns of people in an immediate problematic situation and the goals of social science by joint collaboration within a mutually acceptable ethical framework". A prominent feature of action research is an active and deliberate involvement of the researcher in the context of the investigation. This is unlike many methodologies where the researcher is seen as a spectator who observes different phenomena without intervening.

2.3 Drivers, Inertia and Change

A driver behind large scale changes in organisation can be to keep up with the latest management ideas. This behaviour can be interpreted, according to the new institutional theory (Mayer & Rowan, 1977; DiMaggio & Powell 1983; Czarniawska & Joerges, 1996;, Czarniawska & Savón, 2005), as a need to identify and to legitimize an organisation within the organizational field to which they belong (Czarniawska & Joerges, 1996; Erlingsdottir & Lindberg, 2005). Choices of management methods or processes are hence not necessarily

based on the actual needs of the organization, nor the situation, but are rather a conformity to what the other organizations do and which management methods or processes are in vogue.

Studies of the diffusion of innovations have shown that these diffusion processes can be described by a characteristic "S-curve" (Rogers, 2003). Diffusion is slow in the beginning as only those categorised as "innovators" and "early adopters" are interested in the innovation. But, as the use and knowledge of the innovation spreads, the rate of diffusion increases (fueled also by factors such as maturity of the innovation, lowered cost of providing the innovation, etc.) because the interest of the majority of potential users is won. As the diffusion process approaches saturation, the rate of diffusion decreases, with the group categorised as "laggards" being the only one group (out of those who will eventually adopt the innovation) that has still not adopted the innovation but proceeds to do that at a slow pace. (Rogers, 2003) This theory describes the voluntary adoption of innovation over an extended period of time, and as such is not applicable to the introduction of innovations in a workplace where they subsequently become mandatory to use. However, the fact, that adoption of mandatory innovations is commonly concentrated to a very short interval of time (or even intended to be near instantaneous) makes it interesting to discuss what happens when everyone is thus effectively forced to be an early adopter.

Due to the variation in time and effort required for different organisations to technically and organisationally prepare for adopting an innovation, it is not surprising that the rate of diffusion on a societal level is normally distributed. However, the social factors affecting an individual's voluntary adoption of an innovation usually result in a normal distribution as well, while efforts toward facilitating user adoption strive to shift the weight of the distribution towards more innovators and early adopters while simultaneously maximizing market share (Cain & Mittman, 2002). In sociotechnical systems, the possible rate of change is determined by social, technical, and organisational inertia (Lind, 2014). If these sources of inertia are not taken into consideration during organisational change efforts, the adoption of the innovation is likely to be negatively affected due to the interrelated nature of the organisation of work, the technology supporting that work, and the workers.

3 The Case and the Project

A large Swedish government agency GovOrg is in the process of relocating 600 employees to a new office building that will be ready in 2017. Plans for a new building have been long in the making, but recently management

suggested that in addition to relocating the work force to these new facilities, there should also be a major transformation of work practices and ICT support through the introduction of ABW.

This presented an interesting opportunity for the research group. The research group had previously identified the need for organisational support regarding usability issues within GovOrg (Cajander et. al. 2014). One recommendation was to form a competency centre to promote best practices in the field of HCI and the development and introduction of new competencies and processes for the deployment of IT systems. The research group suggested to the GovOrg board that this concept might be even more relevant to the wider HWID issues related to ABW, e.g. by aiding in identifying how the suggested change could positively affect the employees and how ICT could support the employees in the new work environment. The Swedish Governmental Agency for Innovation Systems (Vinnova) had just started a program for innovation in public organisations and a joint application was made for a planning project. In the end Vinnova granted a planning project entitled Tactics for Activity Based Usability (TABU). The short term goal of the project was to prepare a large proposal for Vinnovas next application cycle in the fall of 2015, with the long term goal of establishing a governance model for sustainable usability in a complex and activity based working environment – to be used in GovOrg but also to be made available to other organisations. It should be noted that the express objective of Vinnova is to support innovation rather than research, though we here use the terms research project/innovation project interchangeably.

The TABU project was composed of separate work packages. The first included business environment analysis and situation analysis and the researchers participated at industry conferences and seminars and also established contact with Swedish researchers in the field as well as with other actors in the area. An important outcome of the activities in this work package is that we have been able to confirm that the project questions are relevant also outside of GovOrg. The group also had access to internal information gathering within GovOrg, such as surveys and meetings. The final work packages concerned the planning of and application writing for the actual innovation project. As there is a gap in time between the end of the planning project and the next grant application cycle, we are using the time to adjust the initial ideas so as to better align with the dynamics of GovOrg.

4 Analysis and Discussion

Results show that the fashionable management approach presented by industry clashed with the vision of many employees in the organisation. There was a considerable gap between the vision for ABW as it was presented by industry and its acceptance by employees at GovOrg. The employees were not swayed, and instead significant social inertia was generated towards the ABW concept and the whole relocation project, which became difficult to handle in the action research project. The pioneers of ABW described it like *the silver bullet*, while many employees were strongly opposed to the idea, in many cases on very rational grounds. This social inertia could be related to the employees experience of GovOrg's technical and organisational inertia (Lind, 2014), and their doubts about the feasibility of such significant organisational changes within such a narrow timeframe as warranted for the introduction of ABW and necessary ICT solutions. Between these two positions, actors on management level were sometimes found to be trying to strike a balance and sometimes to be promoting one side or the other. This setting made the role of the researchers very precarious. As the internal dynamics of the organisation became more sensitive, supporting the TABU project was not a priority on the management agenda and consequently the research group had to accept less insight into the process. In the following we discuss some early and inconclusive observations that we think need to be taken into account when developing the project plan and the project's goal. Using the categorization presented earlier on diffusion of innovations (Rogers, 2003), the industry representatives can be seen as the innovators, the management as the early adopters and the employees take the role of laggards. This is an oversimplification, however, as there are people in the different groups that adhere to other categories.

4.1 The Innovators from Industry

One of the driving forces behind ABW is the available technology, and the large ICT companies are trying to create a new market for communication technology and mobile work technology in this area. When we visited a large ICT company, for example, they first presented their ABW and then they had a presentation of the technology supporting ABW. Conferences on ABW also incorporate presentations of technologies such as Internet of Things, monitoring through sensors and GPS technology for office space. It seems, that ABW is used as a unique selling point for some companies, and to differentiate companies exactly as innovators.

The Early Adopters: Management. Reorganising according to ABW has become a way of legitimizing and branding the organisation. As presented, ABW is as a cool new way of working where everyone is responsible for ensuring that they choose the work place that suit them best given the task at hand. The concept has become a fashion statement, and some of those who were in favour of the concept at the attended industry conferences did indeed present it using the latest marketing strategies. It seems the choice of ABW was based on many things other than the organisation's actual needs, and that at some places visited by the researchers it had the consequence of there simply not being enough room for everyone to attend when they called to large meetings in the organisation.

The Employees as the Laggards: While it should be noted that quite a few employees were positive to the idea of ABW, many of them were very worried and thought that the proposed change was a bad idea. Some argued that it was an idea solely based on economic calculations, as one of the major costs in an organisation is office space. An internal survey the researchers took part of revealed that the majority were indeed very negative to the concept. Many of the employees worried about their work procedures, and how they would be affected when ABW would be introduced, as then their classifiers with papers and all their paper based documentation should be gone.

5 Conclusions

During the planning project it became obvious for the researchers that ABW is an elusive concept. For some people it incorporates management ideas about responsibility and motivation. For others it was connected to the use of technology and for example the ability to monitor whether dustbins are empty in an office. Some seemed to rationalise ABW as a synonym for open office spaces. This confusion regarding the meaning of the concept lead to communication problems during discussions, as people could have so different interpretations.

The challenges presented above lead to dilemmas when managing different roles as a researcher. It might sound easy to work with collaborative action research, and to be a change agent in organisations at the same time as you do research on the change. However, this role is all but simple and requires much experience and skill, as illustrated for example by Westlander (2006).

One of the biggest dilemmas for the research project was that the various positions became entrenched very quickly. In an ideal case the research group might have contributed to finding a common ground between the proponents and opponents of ABW, but in this case there was an obvious risk of the

researchers being perceived as supporting one position or the other. The strong emotions connected to ABW thus caused ethical dilemmas for the research group: How should the research group participate in the discussion without being perceived as taking sides? How normative can we be as researchers when asked about things that are very complex such as ABW? These strong emotions also connected to another dilemma in the project: How much time are we spending trying to understand ABW and its dynamics (ie. reading papers on workplaces, participating in conferences and understanding the concept) in relation to the participation in the planning of the future office building incorporating ABW and handling the internal conflicts in the organisation?

While it would perhaps be easier to change our research focus, we believe that many important design decisions are being made under less than ideal circumstances, such as the ones described above, and that studying them is worthwhile. Still, minor changes to the project idea might make it more feasible. Such changes might relate to the goal as well as the process. Of course, different sources of funding might open new doors (and close others) in terms of alternate directions for the project.

Acknowledgements

The planning project TABU, is financed by Vinnova - the Swedish Governmental Agency for Innovation Systems (2014-03074).

References

[1] Alexander, K. The application of usability concepts in the built environment. Journal of Facilities Management, 4(4), 2006, 262–270.

[2] Appel-Meulenbroek, R., Groenen, P., & Janssen, I. An end-user's perspective on activity-based office concepts. Journal of Corporate Real Estate, 13(2), 2011, 122–135.

[3] Bodin Danielsson, C. The Office–an Explorative Study. Architectural design's impact on health, job satisfaction and well-being. (Doctoral thesis), KTH Royal Institute of Technology, Stockholm, 2010.

[4] Cajander, Å., Nauwerck, G. and Lind, T. Things take time: Establishing usability work in a university context. Proc. EUNIS 2014: Higher education in the digital era, 2014.

[5] Cain, M. & Mittman, R., Diffusion of Innovation in Health Care, Institute for the Future, 2002.

[6] Czarniawska, B. And Joerges, B. Travels and Ideas. In Czarniawska and Sevón eds. Translating Organisational Change. de Gruyter, Berlin, 1996.

[7] Czarniawska B., and Sevón, G. Global Ideas: how ideas, objects and practices travel in the global economy. Liber, Malmö, 2005.
[8] De Been, I., and Beijer, M. The influence of office type on satisfaction and perceived productivity support. Journal of Facilities Management, 12(2), 2014, 142–157.
[9] DiMaggio, P. J., and Powell, W. W. Introduction, in Powell, W. W. & Di Maggio, P. J., eds. The New Institutionalism in Organizational Analysis, The University of Chicago Press, Chicago, 1991.
[10] Erlingsdóttir G., and Lindberg, K. Isomorphism, Isopraxism and Isonymism: Complementary or Competing Processes? in Czarniawska B., & G. Sevón (eds), Global Ideas: how ideas, objects and practices travel in the global economy, 47-70. Liber, Malmö, 2005.
[11] Gates, B. The New World of Work, 2005. Retrieved October 17, 2014, from https://www.microsoft.com/mscorp/execmail/2005/05-19newworldofwork.mspx
[12] Heron, J., and Reason, P. The Practice of Co-operative Inquiry: Research 'with' rather than 'on' people. In P. Reason and Bradbury, H. eds. Handbook of action research, Sage Publications Ltd., 2006, 144-154.
[13] Ianeva, M., Faure, S., Theveniot, J., Ribeyron, F., Cordon, G., Gartiser, C. Pervasive Technologies for Smart Workplaces: A Workplace Efficiency Solution for Office Design and Building Management from an Occupier's Perspective. Proceedings of the HWID conference, 2015.
[14] van Koetsveld, R., and Kamperman, L. How flexible workplace strategies can be made successful at the operational level. Journal of Corporate Real Estate, 1(4), 303–319, 2011.
[15] Laing, A., Duffy, F., and Jaunzens, D. New Environments for Working: The Re-Design of Offices and Environmental Systems for New Ways of Working, 1998.
[16] Lind, T., Change and Resistance to Change in Health Care: Inertia in Sociotechnical Systems, Uppsala University, 2014.
[17] Mayer, J. W., and Rowan, B. Institutionalized Organizations - Formal Structure as Myth and Ceremony. American Journal of Sociology, VOL 83, No 2, 1977, 340-363.
[18] McKay, J., and Marshall, P. The Dual Imperatives of Action Research. Information Technology & People, 14(1), 2001, 46-59.
[19] Oates, B. Researching information systems and computing: Sage Publications Ltd, 2006.
[20] Rasila, H., Rothe, P., and Kerosuo, H. Dimensions of usability assessment in built environments. Journal of Facilities Management, 8(2), 2010, 143–153.
[21] Rapoport, R. Three Dilemmas in Action Research. Human Relations, 23(6), 1970, 499-513.
[22] Rasmussen, L. B. Action research—Scandinavian experiences. AI & Society, 18(1), 2004, 21 - 43.

[23] Reason, P. Choice and quality in action research practice. Journal of Management Inquiry, 15(2), 2006, 187-203.
[24] Rogers, E. M., Diffusion of Innovations, Fifth Edit. New York, NY, US: Free Press, 2003.
[25] Sandblad, Bengt et al. Work environment and computer systems development. Behaviour and Information Technology 22(6) 2003, 375-387.
[26] Too, L., and Harvey, M. "TOXIC" workplaces: the negative interface between the physical and social environments. Journal of Corporate Real Estate, 14(3), 2012, 171–181.
[27] Westlander, G. Researchers roles in Action Research, 2006.

The Form of HWID Theory

Torkil Clemmensen
Copenhagen Business School
Tc.itm@cbs.dk

Abstract. The aim of activities within the Human-Computer Interaction (HCI) area named Human Work Interaction Design (HWID) is to establish relationships between empirical work-domain studies and recent developments in interaction design. Recent areas of research within HWID include design sketches for work, usability in context, work analysis for HCI, and integration of work analysis and interaction design methods for pervasive and smart workplaces. Across these areas, the question emerges what form of theory may HWID research produce? The aim with this paper is to investigate the requirements of different research purposes to a common framework. We take the position that we should approach HWID with a lightweight, medium-level framework that is useful to guide the application of other theories to study the relation between work analysis and interaction design. We analyse the requirements to theory found in research presented at a recent working conference, and find that a frequent form of HWID theory may be a combination of a HCI theory, domain-specific theory and theorizing on the relation between empirical work analysis and interaction design. At the end of the paper, we propose some constraints for further development of a HWID framework.

1 Introduction

The aim of activities within the Human-Computer Interaction (HCI) area named Human Work Interaction Design (HWID) is to establish relationships between empirical work-domain studies and recent developments in interaction design. Recent areas of research within HWID include design sketches for work [1], usability in context [2], work analysis for HCI [3], and integration of work analysis and interaction design methods for pervasive and smart workplaces [4]. Across these areas, the question emerges what forms of theory may HWID produce? The aim with this paper is to investigate the requirements of different research purposes to a common framework. We take the position that we should approach HWID as a lightweight, medium-level framework that is useful to guide the application of more mature theories to study the relation between work analysis and interaction design. Thus, the form of HWID theory will often be a combination of a HCI theory combined with domain-specific theory and focused on the relation between empirical work analysis and interaction design.

The background for doing HWID is that increasingly work environments are challenging for research and design (see the cfp for this workshop). Furthermore, pervasive and smart technologies have pushed work-place

configuration beyond linear logic and physical boundaries. As a result, workers' experience of and access to technology is increasingly pervasive, and their agency constantly reconfigured. While this in certain areas of work is not new (e.g., technology mediation and decision support in air traffic control), more recent developments in other domains such as healthcare (e.g., Augmented Reality in Computer Aided Surgery) have raised challenging issues for HCI researchers and practitioners. The question that such research addresses is how to improve the quality of workers' experience and outputs. The answer to this question will help support professionals, academia, national labs, and industry engaged in human work analysis and interaction design for the workplace. The answer includes developing tools, procedures, and professional competencies for designing human-centred technologies for pervasive and smart workplaces [4].

However, to generate knowledge requires more than results and tools. How to generate knowledge addressing the design of interactive artefacts for challenging workplaces and work environments is a question of theory. The general question of how to theorize about the relations between empirical work analysis and interaction design is the concern of this position paper. In the remaining part of the paper, we analyse the requirements to HWID theory in the different papers from a recent research event (HWID2015). We discuss our findings briefly and indicate some constraints for further theory development. At the workshop, we hope to engage the audience in a discussion about how theories can be adapted, developed, and applied within the research area, and to identify suggestions for further development of the HWID framework.

2 User requirements analysis

Based on the papers, reviews of the papers, and discussions of the paper presentations at the recent HWID2015 working conference, we aimed at identifying the users' (i.e., the researchers') requirements to HWID theory (what theory should be able to do, the form it should have, and what kind of research it would support, etc.). Below we list and explain the requirements that we identified:

1. *Connect Human Work (HW) and Interaction Design (ID)*
2. *Useful for interventions in practical, real-world domains*
3. *Close to an HCI concept, e.g. UX*
4. *Explains aspect of real-world phenomena*
5. *Specify the outcome state of system use*
6. *Deal with context in precise and dynamic way*

7. Multilevel, take organizational aspects into account
8. Describe a work domain
9. Well-known and proven (cited)
10. Socio-technical in the widest sense
11. Useful for domain analysis, eliciting tacit knowledge
12. Can connect a series of domain specific theories
13. Useful to generalize findings to similar domains

1. Connect HW and ID. One requirement to HWID theory is that it can be used to give an overview of research in the field. Frederica Gonçalves, Pedro Campos and Torkil Clemmensen [4] use a HWID framework to present a literature review of 54 papers from HWID workshops, conferences and journals from the period 2009-2014. The paper ends up asking why there is a gap in terms of overall HWID theory development? Questions unanswered in the paper includes if some of the research touch on this partially? Or is HWID theory an integration of theory on three kinds: work analyses, interaction design, and IT artefacts?

2. *Useful for interventions in practical, real-world domains*. Ørngreen [4], in her paper on reflections on design-based research in online educational and competence development projects, is concerned with the usefulness of HWID theory for interventions. She integrates a literature review with her personal experiences in the field to understand the research and practice domain. It turned out from the discussion at the conference that to make her analysis and HWID theory development practically useful, it would be good to have a graphical overview of the different studies presented and how they were related.

3. *Close to an HCI concept*. A third requirement to HWID theory is that it is associated closely with existing HCI concepts or theory, and in this way can contribute to the body of knowledge in HCI. Daniela Wurhofer, Verena Fuchsberger, Thomas Meneweger, Christiane Moser and Manfred Tscheligi [4], in their paper on insights from UX research in the factory on what to consider in interaction design, connect HWID analysis to the HCI concept of UX. Based on recent research (from the last five years), they provide a discussion on the interplay between user, system and context in a factory environment by pointing out relevant UX factors and influences on the workers' experiences. What emerge as a finding is a list of many factors, user experiences and influences, which potentially could be useful for rewriting HCI textbooks on UX.

4. *Explains aspect of real-world phenomena*. This requirement for HWID theory concerns the need to be able to explain interesting aspects of the real-world phenomena studied. Maria Ianeva, Stephanie Faure, Jennifer Theveniot,

François Ribeyron, Gilles Cordon and Claude Gartisier [4] take an industry perspective on pervasive technologies for smart workplaces and in particular a workplace efficiency solution for office design and building management from an occupier's perspective. They focus on how to increase workplace efficiency in the long-term, and contextualize a contemporary concept (Activity Based Workplace) in a living-lab setting (a company's new HQ). They partially explain the success and failures of the implementation studied. Questions to this research include if the specific technological solution for monitoring spaces' occupancy bias the results, and for how long (temporal length) such a study should continue.

5. *Specify outcome state of system use*. This requirement for HWID theory concerns the need for being able to specify the (ideal) outcome state of the analysis and design. Valentin Gattol, Jan Bobeth, Kathrin Röderer, Sebastian Egger and Manfred Tscheligi [4], in their paper on bottom-up insights that leads to design ideas in a case of designing office environments for elderly computer workers, aim to design a smart work environment for this target group. Different research methods were employed including interview, an ideation workshop, and an online survey. Questions raised during discussion concerned what the target situation was for the studied user group. Thus, without supporting theory about human work, it is hard to say what it means to be "old" in contemporary societies (this could include less tech-savvy, myopic vision, back pain, poor fine motor control, memory drop outs, etc.), as well as what kind of interaction design will be useful (e.g., why would a "smart" work environment be considered relevant? How about conventional solutions such as digital reminders and task lists?)

6. *Deal with context in precise and dynamic way*. Väätäjä [4], in her attempt to characterize the context of use in mobile work, provides synthesized findings from twelve cases studies, and derives a model to be used when designing and evaluating systems for mobile work. One question that emerged during discussions of this study was if mobile work context as in this study is best described as a mixture of the context as container (precise description) and context as activity (dynamic description)?

7. *Multilevel, take organizational aspects into account*. This requirement for HWID theory concerns the need to take organizational aspects into account. Petra Bjorndal, Elina Eriksson and Henrik Artman [4], in their attempt to make sense of user-centered perspectives in large technology-intensive companies by looking at relationships rather than transactions, find that theory needs to be multilevel. They studied service design in four IT companies and concluded that studying single human-computer interaction in

a work context (transactions) was far from sufficient to appreciate and design services (relationships) in the organizations.

8. *Describe a work domain*. This requirement for HWID theory concerns the need to describe a work domain. Jose Abdelnour-Nocera, Samia Oussena and Catherine Burns [4], in their study of the Smart University and how to support students' context awareness, created a set of design guidelines, based on the insights from a cognitive work analysis (CWA). As HWID stems from CWA, this is a classic example of HWID analysis (though only in the direction of work analysis towards interaction design). However, the pro and cons of limiting the analysis of relations between work analysis and interaction design to the form of cognitive work analysis (i.e. CWA table) is a topic for future research.

9. *Well-known*. A HWID theory needs to build on the well-known and simple. Åsa Cajander, Marta Larusdottir, Elina Eriksson and Gerolf Nauwerck [4], in their study of IT-based administrative work, brought life into the HWID analysis by developing a concept of "contextual personas". They motivate this in the convergence of interaction design and pervasive workplaces. The main new thing that they introduce to the persona are the statements about control, demand, and support in the work and life of the persona, based on the theory of healthy work by Karasek and Theorell. A question for further development of the concept is how distinct contextual personas are when compared to task and use scenarios?

10. *Socio-technical in the widest sense*. This requirement for HWID theory concerns the needs for dealing with the wider context of use including societal values. Arminda Guerra Lopes [4], in her study of the work and workplaces in social solidarity institutions with the aim to address organization agility and innovation, found that she needed to study the goals and history of the Portuguese social security system. This included that she studied the societal values related to life as an older person. This study uses multiple theories to address the HWID relations and develops a prototype for integration of information related to care services. It provides a clear requirement to the HWID framework to be an organizing framework for multiple theories on many aspects of the context of use.

11. *Useful for domain analysis, eliciting tacit knowledge*. This requirement for HWID theory concerns the need for doing domain analysis and eliciting tacit knowledge. Samantha Quercioli and Paola Amaldi [4], in their grounded theory study of perspectives on automation amongst aviation industry stakeholders, learned that systems are far too complex for people to understand. Hence the main target for the domain analysis emerges as trying to identify what half-baked knowledge that human actors need to have about

non-human actors in order to avoid or recover from unsafe situations with aircrafts. It turned out in the discussions at the working conference that in this case human work and interaction design were clearly mutually constitutive.

12. C*onnect a series of domain specific theories*. This requirement for HWID theory concerns the need for connecting several domain specific theories. Judith Molka-Danielsen, Mikhail Fominykh, David Swapp and Anthony Steed [4], in their design of a virtual learning environment to teach space syntax from the user's perspective, created an artefact by combining insights from several domain-specific theories. The theories were a 'threshold' concept from didactics, 'line of sight' from architecture, Virtual Learning Environments (VLE) theory, and more.

13. Useful to generalize findings to similar domains. This requirement for HWID theory concerns the need to make results useful for other similar domains as the one studied. Stefano Valtolina, Barbara Rita Barricelli, Davide Gadia, Matilde Marzullo, Claudia Piazzi and Andrea Garzulino [4] integrates multiple studies of co-design of cloud services for archaeological practice, and identify two problems in current archaeological practice: 1) how to integrate and create useful knowledge from a richness of documentation, and 2) how to facilitate collaboration among various domain experts. They adopt a semiotic approach in combination with tool design to do the HWID. Future research may map the relations between the two theories in order to generalize to other domains and other archaeological sites.

3 Discussion and Conclusion

In this position paper, we have illustrated that a frequent form of HWID theory is as a meta framework that serves to focus discussion on Human Work (HW) and Interaction Design (ID) relations in wide contexts. In the light of the user requirement analysis presented above, it appears that one of HWID's main heritages from cognitive work analysis (CWA) is the need to adapt, combine, and develop a mix of theories in order to understand the relations between HW and ID on multiple levels. Further development of the HWID framework may make it a better tool for mapping the combinations of multiple other theories onto the relations between HW and ID. In addition, the IT artefacts, e.g. contextual personas by Cajander et al (see above) or the threshold artefacts by Moltke (see above), may be examples of hybrid work analysis and interaction design artefacts coming out of HWID. Finally, the HWID theory should reflect that artefacts may be 'packed' with values from both HW and from ID (see Lopes, above).

References

[1] Clemmensen, T., Campos, P., Ørngreen, R., Pejtersen, A. M. and Wong, W., (Eds). Human Work Interaction Design - Designing for Human Work. First IFIP WG 13.6 Conference, HWID2006, February 13-15, 2006, Madeira, Portugal. Revised Selected Papers. Springer, 2006.

[2] Katre, D., Orngreen, R., Yammiyavar, P. and Clemmensen, T., (Eds). Human Work Interaction Design: Usability in Social, Cultural and Organizational Contexts: Second IFIP WG 13.6 Conference, HWID2009, Pune, India, October 7-8, 2009, Revised Selected Papers. Springer, 2010.

[3] Campos, P., Clemmensen, T., Abdelnour-Nocera, J., Katre, D., Lopes, A. and Ørngreen, R., (Eds). Human Work Interaction Design. Work Analysis and HCI: Third IFIP 13.6 Working Conference, HWID2012, Copenhagen, Denmark, December 5-6, 2012, Revised Selected Papers. Springer, 2013.

[4] Abdelnour-Nocera, J., Clemmensen, T., Campos, P., Lopes, A. and Barricelli, B., (Eds). Human Work Interaction Design for Pervasive and Smart Workplaces. HWID2015, London, June 25-26, 2015. (In preparation).

Understanding UI Design for Creative Writing: A Pilot Evaluation

Frederica Gonçalves[1], Pedro Campos[1], Anant Garg[2]
[1]Madeira-ITI, University of Madeira
[2]HaritaDhara Research Development and Education Foundation
frederica.goncalves@m-iti.org; pcampos@uma.pt; anantg@hrdef.org

Abstract. This paper describes a new study about how youths react when they use creative writing environments to express themselves. In this pilot study, we present reactions of young students in India, in a between-subjects design using MS Word as baseline and a "calm" and relaxing writing tool called OmmWriter. We analyze the influence of this type of tools to assess if a creative writing user interface can positively influence the productivity and mental well-being of users. Writing has been one of the most important developments of human civilization, but people are often unconscious of how long and complicated this path has been. However, understanding the creative writing user interface is a necessary activity if one wishes to shape future tools. Following a triangulation from results and qualitative data, we could notice that this experience increased the students' desire to write more, every day.

Keywords: Creative Writing tools; Human Computer Interaction; Creativity Support Tools; User Interface Design; User Experience Design.

1 Introduction

The best way to view the proliferation of human civilization is to look at the different ways we have to write. Writing is one of the most important human activities and one of the oldest. The increasing adoption of computerized tools caused people to express themselves more easily and more often.

In this paper we report a pilot evaluation with students in India, using two different creative writing tools: MS Word, our baseline, and OmmWriter a Zen-like tool. We briefly describe the methodology, procedure and some results, triangulating data from different sources to assess if a creative writing UI can positively influence the productivity, mental well-being and creativity of users. We close this pilot study by performing a reflection and discussion about results and also presenting future work.

2 Related Work

Technology is changing over the past last decades. The intensification of digital computing accelerated writing in all its forms, and people in today's society have a much easier access to reading and writing. Whether you're script-writing a novel or imagining a new commercial ad, the process of writing depends on prewriting that leads you after to a draft, which is then to be revised [7]. Chang et al. [3] refer that a good storyteller usually needs good inspiration during the story construction process. Other researchers [4] have studied the role of storytelling technologies to encourage collaboration and to reflect design suggestions made by children themselves. Creativity includes discovery or invention of a significant idea, pattern, method, or device that gains recognition from accepted leaders in a field [5] and is the ability to produce work that is unexpected, high in quality and useful [1]. Some researchers, such as Carrol and Latulipe [2] considered that is a challenge to measure when a person is actually being creative. Broadly speaking, the same authors, argue that there is no consensus in how to measure creativity or when a person is "in the moment" of creativity, and there is a challenge in evaluating creativity support tools. Even with decades of creativity research, there is no single, agreed upon methodology for evaluating how well a creativity support tools to aid the creativity of its users [8]. There are also different studies to overview and measure the creativity behind ideas [9] but a lack of research work in creative writing user interfaces and tools, a gap this paper attends to address by giving insights to future work. Christiaans [10] suggest that as long as no absolute criterion of creativity exists, the assessment of creativity remains dependent on subjective judgment. He refers that design can include more objective aspects that mainly involve the functionality and technical quality of the design.

3 Field Study of Creative Writing User Interfaces

We designed and conducted one pilot experiment in order to investigate if a creative writing UI can positively influence the productivity, mental well-being and the creativity of users. As an introduction to the following experiments, we will focus on some characteristics of participants. Since we were interested in empowering youths by providing them with creative writing tools, we targeted teenagers that normally attend to a non-profit organization in India. The institution develops capacities for building environments in schools for children, quality of life through education, and for students to be capable of having better curriculums and new skills of development. These

teenagers came from middle class families, some already have business and others are working with their parents. Figure 1 illustrates the environment where the study took place.

Figure 1. Youths in the institution establishment.

In the next sections, we will describe the evaluations, including the participants, method, procedure and results of each. All the data taken from the experiments was made completely anonymous.

Study: OmmWriter vs. Word

We conducted a user study to address the following research questions:
RQ1: *can a "Zen-like" UI positively influence the productivity of users?*
RQ2: *can a "Zen-like" UI positively influence the mental well-being of users?*
RQ3: *can a "Zen-like UI positively influence the creativity of users?*

Participants

A total of 10 students were involved in this preliminary study (aged between 10 and 16 years old). There were three females and seven males. We conducted five sessions and used PCs. Participants had computer and Internet experience.

Method & Setting

This study was organized based as a true experiment on a between-subjects setting using repeated measuring. To minimize the existence of confounding variables, two groups of student were created: a control group that used MS Word as a baseline, and an experimental group that used the Zen-like creative writing user interface of OmmWriter. The order of the two conditions was counterbalanced, participants were random allocated and every participant was equally likely to be allocated to each group. Creativity and mental well-being were measured using subjective, quantitative (Liker-based scale) daily surveys which were constructed from the Flow Theory's concepts [6]. Before starting the first session, the experimenter explained the scope of the study and the session rules. Each group was instructed to start by answering the daily

challenges and then responding to a daily survey. During the experiment, the two groups were seated in separated in different rooms.

Procedure

The study consisted in three phases: *writing a specific challenge* using the addressed tool, *answering a survey,* and *rating the creativity* of written data. The time was limited (30minutes) for each writing challenge. No time limit was set for completing the writing tasks. After each session, the participants were asked to fill out a very short Likert scale survey about how the daily experience made them feel.

Writing a specific challenge. In each session, participants received a writing task (writing prompt) to initiate their writing. We gave the same writing challenges to all students. Examples of the writing task included: "*Write about a time when you used your inner strength to get through a tough situation"; "Who is your Hero and why?"; "Write about when someone hurt your feelings"; "Have you ever took a risk?"; "If you could change the world what would you do and why?"*. These were conceived by one of the authors and it was taking into account the age of the participants.

Answering a survey. In each session, participants fill out an online survey that was based on the Flow Theory dimensions: (i) intense and focused concentration on the present moment, (ii) sense of personal control or agency over the situation or activity, (iii) loss of reflective self-consciousness, and (iv) distortion of temporal experience. Participants ranked a seven-point Likert with the evidence scale for 1 (totally disagree) and 7 for (totally agree), based on questions such as: e.g. "*I felt very concentrated during the challenge"; "I lost track of time during the challenge"*, etc.

Rating the creativity. After finishing the five sessions of the experiment, we asked participants to read their written data, and to rate their creativity. They were asked to provide a value ranked in a 5-point Likert with the evidence scale for 1 (not really) and 5 (very much).

Results

We evaluated the study from a perspective that triangulates the results, using the answers from the surveys, the statements of the interviews and the qualitative measurable information of the writing challenges. Results show that the number of test items can be considered with a good consistent in the scale used from the questionnaire on seven-point Likert scales. From our control group and experimental group we can view in Table 1 the listed demographic characteristics for each group.

		Control Group	*Experimental Group*
Age	10-11	1	2
	12-13	2	2
	14-15	1	1
	16	1	0
Gender	Female	1	2
	Male	4	3
	No. Participants	5	5

Table 1. Demographic information for participants in each group

To assess the productivity of the users we started counting the number of words written per participant in each tool (Table 2). Regarding data dispersion users in the control group using MS Word (baseline) wrote more words, than those in the experimental group using OmmWriter.

	Microsoft Word	*OmmWriter*
Mean	45.80	42.08
St. Error	6.7	5.1
Median	34	35
St. Deviation	33.66	25.44

Table 2. Statistics from number of words written

Although students were curious about the experience and excited to know what challenges were proposed, triangulating these results with semi-structured interviews allows to say that there is a tendency to increase their desire to write more, and we noticed that day by day. Note that the challenges were the same for both the control group and experimental group, and there was no transfer of learning as we were using a between-subjects experimental design. None of the users had ever taken an experiment such as this one. We could notice that some students liked to answer some challenges than others. When we asked them why that liked, they argue that writing tasks make them thought more about, and more things came into their minds. All participants in this experience, had experience with MS Word, but had no experience with OmmWriter. As we can see from the results, the experimental group showed similar results as the control group, despite the fact that they had never worked with that tool before. To access the participants' mental well-being as well as their evolution along the five days of this study, we asked them to select up to three adjectives from the following list: Surprised, Delighted, Laid back, Depressed, Pacific, Happy, Tired, Bored, Sad, Satisfied, Frustrated, Angry, Serious, Animated, Distressed, Creative and Frightened. Figure 3 displays the total count for each adjective, as selected by the participants. We can see that Happy, Delighted, Satisfied, Relaxed and Serious were the most chosen adjectives, especially when using OmmWriter.

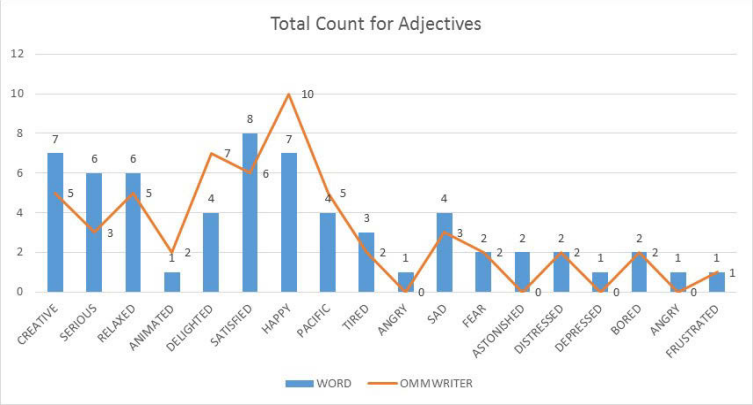


Figure 2. Total count for the adjectives chosen by participants.

By looking at the Figure 3, one can see that the sense of Concentration was not a significant issue for any of the tools we evaluated. The same observation is valid for the Sense of Control dimension, despite the fact that there is a minor difference between tools. The greater difference, was found in Lost of Self-consciousness and Lost Track of Time. According to the interviews, students reported that OmmWriter was effective user interface for feeling better (with music), happier and also more relaxed.

Figure 3. Combined results for five days experience and the different tools involved with Flow Theory.

In the third phase of this experience, we asked students to rate their creativity of all written data per writing challenges (Table 3). This was made after the day five.

	Microsoft Word	*OmmWriter*
Mean	3.32	3.28
St. Error	.21	.20
Median	3	3
St. Deviation	1.07	1.02

Table 3. Statistics about self-rating creativity from written data

Regarding the stories written, and from a creativity perspective, we found some differences between the control group and experimental group, with participants using the calm and relax tool apparently allowing for greater levels of creativity. However, it was not possible to establish this difference with sound confidence. When participants in the experimental group

(OmmWriter) were interviewed, they especially said they liked the background soundtrack, because helped them to express more and in a quiet way, e.g. "*the continuous soothing music, help me to express…"-userOmmWriter.* Based in our observations and from the written data, it can be said that experimental group have a tendency to perform more in future experiences. When interviewed, all express genuine interest in creative writing and most users emphasized that wanted to do this experience for more days, using the tools and the writing tasks *"…it's a good started for writing…"-userMSWord; "I was curious to find out each day, what was supposed to write…"-userOmmWriter.*

4 Discussion and Conclusion

The scope of the pilot study discussed in this paper was design to investigate if creative writing environments can encourage youths to express themselves through creative writing. It was limited to a small number of participants in order to develop analytic methods with a data set. Although the study involved students in India attending the institution, the methods proposed in this study could be used in future work in different institutions. To further shape an evidence base, the study needs to expand to include more days and more different creative writing environments. Also, we need to investigate different ways to measure the creativity, and how it is maintained, and for how long during the long-term usage of the tools. Despite the fact that the pilot study included only a small number of collected data, the need for substantive instructions and encouragement for creative writing environments being used by novices was strongly suggested. During all experiments, it was clear that participants became intensely concentrated to solve the writing challenge. From a qualitative perspective, a loss of reflective self-consciousness was also reported, especially using the OmmWriter tool. Also, it was clear that every participant considered that the writing tasks were a good kickstarter for the creative writing process. We could notice that participants felt somewhat empowered and creative during the experiments. Future work will compare the impact of creativity, satisfaction, productivity per daily challenges in students who have different access to education and will examine the use of different user interfaces for creative writing can empower to express themselves and give them a voice. The biggest limitation of our study it that it doesn't consider the long-term usage of these tools. Therefore, conclusions are limited to an incipient (five days) usage of the different creative writing user interfaces.

References

[1] Sternberg, R.J., Lubart. T. I., Kaufman, J. C., Pretz, J. E. (2005) Creativity. In K. J. Holyoak & R. G. Morrison (Eds.), The Cambridge handbook ok thinking and reasoning, pp. 351-369.NY: Cambridge University Press.

[2] Carrol, E.A., Latulipe, C., (2011). Capturing "In the Moment" Creativity Trough Data Triangulation. *C&C'11*, Atlanta, Georgia, USA.

[3] Chang, Y., Chen, S., Li, T. (2011). A Computer System Aiming to Stimulate Creativity for Narrative. *C&C'11*, Atlanta, Georgia, USA.

[4] Benford, S., (Eds.) (2000). Designing Storytelling Technologies to Encourage Collaboration Between Young Children. Proceedings of CHI2000.

[5] Schneiderman, B. (2007). Accelerating Discovery and Innovation. *Communications of the ACM, 50, 12*.

[6] Csikszentmihalyi, M. (1990), *Flow: The Psychology of Optimal Experience*. New York, NY: Harper and Row.

[7] Campos, P., F. Gonçalves, M. Martins, M. Campos, and P. Freitas (2014). "*Second Look: Combining Wearable Computing and Crowdsourcing to Support Creative Writing*", Proceedings of the 8th Nordic Conference on HCI. Helsinki, ACM, pp. 959–962, 10/2014.

[8] Erin A. Carroll and Celine Latulipe (2012). Triangulating the personal creative experience: self-report, external judgments, and physiology. In *Proceedings of Graphics Interface 2012* (GI '12). Canadian Information Processing Society, Canada, 53-60.

[9] Silva, P.A., Read, J.C. (2010). A Methodology to Evaluate Creative Design Methods: A Study with the BadIdeas Method. OZCHI 2010,Brisbane, Australia.

[10] Christiaans, H. C. M. (2002). Creativity as a Design Criterion. Creativity Research Journal 14(1), 41-54.

Pen and Display: A Multimodal Interaction Approach for Older Office Employees

Georg Regal[1], Ulrich Lehner[1], Valentin Gattol[1], Jan Bobeth[1], Manfred Tscheligi[1,2]

[1] AIT Austrian Institute of Technology GmbH
Innovation Systems Department, Technology Experience
Giefinggasse 2, 1210 Vienna, Austria
[firstname.lastname]@ait.ac.at

[2] University of Salzburg
Center for Human-Computer Interaction
Sigmund Haffner Gasse 18
5020 Salzburg, Austria
manfred.tscheligi@sbg.ac.at

Abstract. Pen and paper are still common tools in the daily work routine, especially of older office employees. In this position paper we present "Organio", a concept of combining (analogue) pen and paper interaction and page tracking with a digital display. This novel digitally-enhanced pen and paper solution relies on a page tracking mechanism to provide digital content to the selected page in an analogue calendar. Building on a recent case study research on older computer workers and related work we describe a first calendar mock-up enhanced with a digital display, an interaction concept, and a setup for evaluating the proposed solution in a *Wizard of Oz* study. We conclude with ideas for future study concepts to investigate the overall user experience, user acceptance and possible form factors. Moreover using the *Wizard of Oz* approach we aim for gathering user expectations and possible usage scenarios.

1 Introduction

The work routine of older office employees still relies heavily on pen and paper, as observed by Gattol et al. [4] in a recent case study. In this position paper we investigate possibilities to digitally support older office workers without changing the pattern of analogue pen and paper interaction.

We focus on how to support older office employees with a mixed modality interaction approach that capitalizes on interaction patterns familiar to this target group, yet compatible with today's digital office environments. We present a concept of combining analogue pen and paper interaction with a digital display to allow for multimodal interaction, called *Organio*. We address common time and task management needs and explore possibilities to integrate familiar analogue forms of interaction (e.g. note taking) into digital work processes (e.g. meeting scheduling).

2 Related Work

The aim of this position paper is to describe a concept that enhances the current workflow of older adults. In [4] Gattol et al. observed that analogue calendars are still a main source to store all kinds of information, e.g. appointments, to-do lists, etc. Yet, this behavior can be cumbersome, keeping in mind that information with colleagues and customers is often exchanged by email and meetings are scheduled electronically in a groupware system. Presently, this results in redundant (and thus often inconsistent) storage of information, e.g., when a meeting is scheduled electronically and then written down in the analogue calendar as well.

To provide appropriate digital content to a certain calendar page a method to track the page of the analogue calendar viewed by the user is needed, to provide the corresponding information on the digital display. So far only few approaches for tracking pages have been proposed. Iwaki et al. [7] proposed tags on paper and a camera system to track and identify pages. Back et al. [1] used RFID tags embedded in the pages. Fujinami et al. [3] used accelerometer data for detecting page flipping events. For our work we will follow the approach by Back et al. [1], as this seems to be the most accurate solution.

Multiple approaches have been proposed to support pen and paper interaction. For example Heinrichs et al. proposed design recommendations and a meta model for pen paper interaction [5,6]. Chuang et al. [2] have presented a system to support students by enhancing textbooks with a digital pen. Related to our approach is the work of Williamson et al. [8] who have proposed a pen and paper based reminder/calendar system in the domain of Ambient Assisted Living (AAL). In contrast to our approach, events written down in a paper calendar have only been used to trigger reminders but have not been integrated into a more complex routine like office work. Apart from that our contribution is also the strong focus on the context of office work and on older employees, which results in different design considerations than reminder systems in the domain of AAL.

3 Development of the Concept

Our main research question is how we can support the work routine of older computer workers that relies heavily on pen and paper interactions.

We introduce a concept for a digitally-enhanced pen and paper solution that can help cross the digital chasm and integrates into the work routine of older office employees.

This idea is based on insights derived from a case study research on older computer workers, conducted in a company developing and maintaining point of sale payment systems. The case study relied on a multi-method approach (i.e., contextual interviews & observations, an analysis of needs and frictions, a feature ideation workshop, and an expert evaluation of the resulting feature ideas) for specifying the user requirements [4]. The following three insights, reported in the above paper and formulated from the perspective of the user, give a glimpse of the underlying need in our target group for some form of traditional or analogue time and task management solution that is also compatible with today's digital office environments (pp. 6):

- "I need to keep track of open tasks but carrying them over (e.g., from my agenda to Outlook) takes effort."
- "I need to share information digitally but there is no efficient way of digitizing hand-written information."
- "I plan my day on paper because I don't see the benefit of digital solutions."

Based on these insights the authors reported the following preliminary feature ideas (pp. 7):

- Private Digital Noteboard: an always visible second screen at personal desk, used for highlighting urgent tasks, clustering tasks, etc.
- Public Digital Noteboard: an always visible second screen at a wall, used in conjunction with the Private Digital Noteboard.
- Digital Paper Calendar: a paper calendar capable of automatically digitizing hand-written notes, using digital paper or a digital pen as input.

Thus, given these insights and preliminary feature ideas, our idea for a novel digitally-enhanced pen and paper solution addresses a common and relevant need of older office employees (cf. [4]). In our work we combine page tracking and pen and paper calendar systems to develop a solution specifically for older office employees to be used in their normal work environments. Although parts of our approach have been covered in related work [1,8], to the best of our knowledge no similar system has been proposed in the context of work support for older office employees. Therefore we enhance the work of Williamson et al. [8] by providing a digital layer of information to every calendar page by using page tracking approaches by Back et al. [1].

In the following two sections we will elaborate on the interaction concept that we conceived of and describe a setup for evaluating the prototype in a Wizard of Oz study.

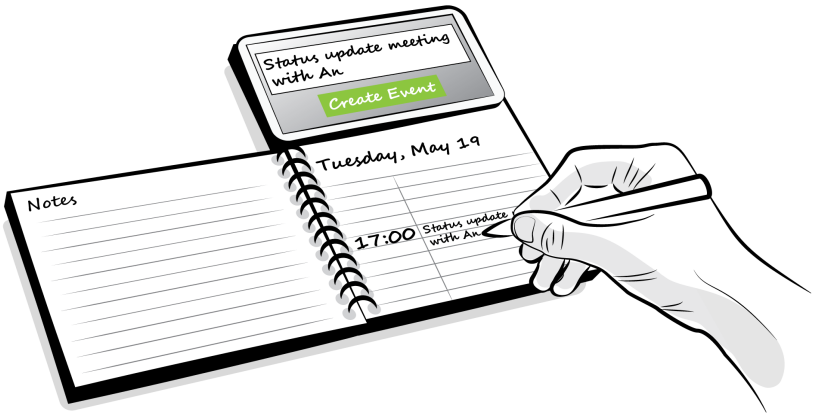


Figure 1. Mock-up of the Organio prototype concept - e.g. an employee is writing a calendar event which is previewed on the display and suggests creating an event

3.1 Interaction Concept

Organio aims to fit into typical workplace use cases, such as calendar applications (e.g. to organize meetings) or note taking (e.g. to manage a list of to-dos). The proposed concept, as illustrated in Figure 1, consists of a paper calendar with a touch display on top that is aware of the current page. Furthermore, it is supplemented with a smart pen which digitizes the hand-written input and tracks the position on the paper. The pages are prepopulated with templates to fill in notes (left page) and calendar events (right page).

The combination of these technologies enables a variety of interactions we want to investigate:

- *Touch input:* Touch input on the display allows recognizing single-touch (e.g. tapping a button) or multi-touch (e.g. pinch gesture to zoom) gestures.
- *Graphic output:* A graphic display for displaying digital content (e.g. event invitations).
- *Digitized hand-writing:* Analogue hand-written content is digitized by a smart pen and enables OCR analysis or tracking the pen position on the paper (e.g. to detect the time of the day).
- *Page flipping:* An essential interaction with *Organio* is the page-flipping mechanism that allows viewing content corresponding to a specific page (e.g. shared calendars of colleagues based on the date of the current calendar page).
- *Pen gestures:* Utilizing the smart pen for interactions with the written content on the paper opens several possibilities (e.g. mapping common hand-writing behaviors to certain actions: striking through words to delete a to-do; writing a check mark next to a line to mark a to-do as done). These interactions are common routines and can be synchronized

with a digital system. Other gestures can include circling words to select them (although without using real ink) or introducing "action words", e.g. a "T" as a prefix to indicate that the following content should be interpreted as a to-do.

The proposed combination of page tracking, pen and paper input and a digital display strives to seamlessly integrate common behaviors with widely used paper-based calendars by focusing on interaction patterns and habits that are familiar to our target group of older office workers. For example by using *Organio* a user can take notes on paper for preparing a meeting, while at the same time the user can participate in an organization's digital workflow (e.g. syncing the meetings calendar events with the groupware software). Moreover the user can share the pen and paper notes with his or her colleagues.

3.2 Wizard of Oz Prototype

We propose a prototype that allows investigation of the page tracking mechanism, as this is a crucial part in our concept. We developed a low fidelity prototype, enhanced by interactivity using a Wizard of Oz approach to simulate the page flipping / tracking mechanism for the users. In this prototype we combined a Nexus 5 smartphone and a paper block. To simulate the page flipping mechanism, we developed a smartphone application that allows for remotely switching screens on the *Organio* screen. For this reason two phones are connected using a basic Bluetooth message exchange application. When the user switches a page the facilitator presses a button on the second device, which results in a message being transferred via Bluetooth to the *Organio* and subsequently a reaction based to that message. For example (1) the user flips a page to the left, (2) the facilitator presses a button to change content; (3) a message is sent via Bluetooth from the facilitator's device to the *Organio* and (4) the content appears on the *Organio* screen accordingly.

The described prototype will be used in future studies to evaluate the proposed interaction concept with special focus on the user acceptance of the page tracking mechanism.

4 Discussion and Future Work

In this position paper we proposed a concept called *Organio* for integrating an analogue pen and paper solution into the digital work routine of older office employees. *Organio* provides digital content corresponding to the current page in an analogue calendar and suitable interfaces for sharing and distributing pen

produced content. In the future we see *Organio* in combination with a personal display located on the desk and a public display to share content.

In future work we want to investigate the feasibility and user experience of the overall approach in a co-design workshop and a user study. In the co-design workshop users will be provided with a lo-fi mock-up of the proposed pen, paper and display solution. In the workshop we want to discuss the overall concept, its applicability to different work domains, what kind of analogue content the users are willing to share and,investigate the optimal form factor. Subsequently we will conduct a user study with a Wizard of Oz prototype to simulate the page tracking mechanism for users. Using this Wizard of Oz prototype we aim to investigate the acceptance of the pen, paper and display solution and moreover gain deeper insights into the users anticipation how the proposed solutions fits into their work routine.

Acknowledgements

This work has been partly funded by the AAL JP project PEARL (AAL-2013-6-091).

References

[1] Back, M.J. and Cohen, J. Page detection using embedded tags. *Proceedings of the ACM symposium on User interface software and technology (UIST '00) 2*, (2000), 159–160.

[2] Chuang, C.-H., Chao, P.-Y., Wu, H.-K., and Chen, G.-D. Integrated Textbook: Augmenting Paper Textbooks with Digital Learning Support Using Digital Pens. *Sixth IEEE International Conference on Advanced Learning Technologies (ICALT'06)*, (2006).

[3] Fujinami, K. and Inagawa, N. Page-flipping detection and information presentation for implicit interaction with a book. *International Journal of Multimedia and Ubiquitous Engineering 4*, 3 (2009), 93–112.

[4] Gattol, V., Bobeth, J., Röderer, K., Egger, S., and Tscheligi, M. From bottom-up insights to feature ideas: a case study into the office environments of older computer workers. *Proceedings of the Human Work Interaction Design (HWID) Conference*, (2015).

[5] Heinrichs, F., Schreiber, D., Huber, J., and Mühlhauser, M. W 5 : A Meta-Model for Pen-and-Paper Interaction. *The 3rd ACM SIGCHI Symposium on Engineering Interactive Computing Systems (EICS´11)*, (2011), 47–52.

[6] Heinrichs, F., Steimle, J., Schreiber, D., and Mühlhäuser, M. Letras. *The 2nd ACM SIGCHI Symposium on Engineering Interactive Computing Systems (EICS '10)*, (2010), 193.

[7] Iwaki, S., Kawata, H., Muto, S., Machino, T., Motegi, M., and Nanjo, Y. FreeTagpaper : A Pen-and-Paper-Based Collaboration System Using Visual Tags Printed on Paper x. *International Conference on Soft Computing as Transdisciplinary Science and Technology (CSTST´08)*, (2008), 1–5.

[8] Williamson, J.R., McGee-Lennon, M., Freeman, E., and Brewster, S. Designing a smartpen reminder system for older adults. *CHI "13 Extended Abstracts on Human Factors in Computing Systems on - CHI EA"13*, (2013), 73.

Out in the Cold, the Loneliness of Working with Doctors and Patients

Bert Vandenberghe, David Geerts
Centre for User Experience Research (CUO)
iMinds – KU Leuven
Leuven, Belgium
{firstname.lastname}@soc.kuleuven.be

Abstract. Working as an HCI researcher in the domain of healthcare can be challenging. The nature of the domain limits our degrees of freedom, which we need to do research of high quality. Doctors never have time, while working with them can be time consuming with little tangible outcomes. Patients from their side might bring the researcher in uncomfortable situations, adding little contribution to the research itself but leading to psychological stress and mental burden on the researcher. And as researcher, it is often difficult to get the research published to the full potential due to the limitations inherent to the domain of healthcare. We found sharing experiences very useful when working out solutions for these issues, so we would like to continue this discussion during the workshop.

Keywords: Ethics, Fieldwork, HCI Researcher, Healthcare.

1 Introduction

"A hospital is a closed environment, you don't just walk in here and do whatever suits you" answered the physician in a threatening tone to the question whether I could do a contextual inquiry at 'his' hospital department. As the physician argued, the hospital is a place where patients are being treated, and there is little space for experimentation of someone without a white coat. Patients should feel at ease here, and should be able to trust their physicians. For this reason, we submit our protocols to an ethical committee, who can approve our study from an ethical point of view. However, I find these ethical concerns too unidirectional. The subject of the study, e.g. the patient, is well considered before getting approval. But the other side, the researcher, is often left out in the cold.

After two years of working as a HCI researcher in the domain of healthcare, I experienced a number of challenging situations and learned to deal with them as well. This workshop paper discusses some difficulties when working with doctors and patients. As I found sharing experiences with colleagues or reading case studies (e.g. [1], [2]) the most useful way to tackle these challenges, I would like to continue this discussion during the workshop.

2 Three sides to the story

Doctors and patients have a strong but delicate relationship. Patients literally put their life in the hands of the doctor. The patient needs to trust the doctor in order to make treatment work. Also in our work as HCI researcher, trust and reciprocity are important in the relationship with the users. So when working in the domain of healthcare, the HCI researcher is dealing with doctors and patients, and thus a triangular relationship arises. While confronting doctors and patients can be interesting, we must be careful not to interfere with the doctor-patient relationship, and thus the medical outcomes. On the other hand, we must try not to be the fifth wheel in our own research project. In the following paragraphs, I discuss some challenges of HCI research in healthcare from the perspective of the doctor, the patient, and the researcher.

2.1 Doctors never have time

Doctors, especially specialists, have little time. In Belgium, physicians are paid on performance, which leaves little time for extras. Arranging a meeting can be time consuming, and even during an encounter physicians might need to leave you alone. After all, emergency cases are common in a hospital.

Once, I invited 12 specialists that collaborated in a project for a workshop. Most of them declined the invitation, as their schedule was too busy already. A few did not make it to the workshop because of emergency cases, they said afterwards, and one was late for the same reason. One specialist was on time. Afterwards, the two specialists who did attend our workshop asked us to repeat this exercise with the other specialists because it made them think in a way they were not familiar with. We never did, as we never found a suitable moment.

As their time is limited, working with doctors requires you to be very efficient. On the other hand, it is common for a specialist not to answer my email for weeks and then reply “come and meet me tomorrow”. So in contrast to the requirement of being efficient with their time, I found working with doctors very inefficient from a time management perspective. When working with doctors and specialists, I learned to prepare myself to the maximum and always be flexible with my schedule. I always have a plan B (e.g. some reading material), just in case plans change.

2.2 Patients have uncomfortable stories

Patients, for example chronic patients, often have long stories. A disease never comes alone, and the contextual causes and consequences come with it.

Once, I did a diary study with transplant patients. We visited the patients two times in their home environment, first to explain the study and then at the end for a contextual interview. As required, we submitted our protocol to an ethical committee for review. We had to argue why it was necessary to meet patients in their home environment, as we entered the personal sphere of the patient, which could be too obtrusive for some our participants. We got approval after some discussion and could start the study.

Many, if not all, patients warmly welcomed us. At most places, we received coffee and cookies. We received a lot of valuable input for our study and most participants explicitly stated that they were open for future studies. I was surprised how open people were about their health, and their life.

In this openness, we also encountered all kinds of uncomfortable situations. The participants seemed to be glad that they could share their story, as someone was listening. Stories about relationship problems, divorce, alcohol abuse, depression, and loneliness were told, often in tears. As researcher, I felt mental burden and didn't really know what to do with it. But because we were also gaining a lot of insights, I didn't want to change our approach.

In the same study, we had a patient who had acute problems while completing his diary and he was admitted to the hospital. On a Sunday, while I was preparing the home visits for the next day, I received a voicemail from the hospital. The patient, who sounded very weak and had difficulties to find breath, told me he wouldn't be able to finish the diary. He called me to say he gave his diary to a nurse, so they knew where it was. He apologized because he couldn't finish it as planned, and said goodbye. I couldn't get his voice out of my head for several days.

Before the study, I was afraid of confronting patients with uncomfortable thoughts. However, I experienced the exact opposite and maybe came too close to the participants. As researchers, we are interested in the context in which technology is used. But by scratching the surface, patients might bring us in uncomfortable situations. I learned to schedule enough time when visiting patients. If this buffer is not used, it allows me to clear my head before meeting the next patient by writing down my impressions of the visit.

2.3 Researchers must publish

Working as HCI researcher in the domain of healthcare is not always straightforward. Of course, some of these challenges are inherent to the work as HCI researcher, or to research in general. Doctors don't have time, and asking too much time of patients can be too exhausting for them. So we do get less time than we would like with less doctors and patients, our users. Often,

outsiders are quite surprised that we only worked with 3 hospitals. Others, who are more familiar with the domain acknowledge the difficulties in getting 3 hospitals prepared to open doors to let a stranger in and make time for him. Clinical trials, which can show whether there is an effect take very long, also to get them organized. As we are in a fast-paced domain, these studies can take too long.

Also, in the healthcare domain, we see and hear all kinds of things that we shouldn't. In most situations, taking notes is all we can do. So we also have to be creative on the data gathering and analysis side of our research, which could also be seen as limited by outsiders. While we have the tacit knowledge because we immersed ourselves in the context, we can't back up these insights to the full potential.

3 Conclusion

As HCI researcher, we translate domain knowledge, needs, and opportunities from different points of view. These insights might be obvious for some, but eye opening for others. Especially in the domain of healthcare, being disruptive can be very counterproductive. Doctors are sometimes considered as conservative by nature, and reluctant to change. But in our experience, they care about their patients and their treatment. So we had very interesting discussions with them.

Working in the domain of healthcare challenges us to be creative. By sharing stories, we get the inspiration to tackle obstacles along the road when working with doctors and patients. In this workshop, we would like to discuss how we could adapt our methods to fit these specific situations. How can we deal with the contextual factors of working with doctors and patients? How can we do research that matters, balancing empathy and distance between us and patients or topics? And how can we get our results published?

References

[1] Furniss, D., O'Kane, A. A., Randell, R., Taneva, S., Mentis, H., & Blandford, A. (2014). Fieldwork for Healthcare: Case Studies Investigating Human Factors in Computing Systems. Synthesis Lectures on Assistive, Rehabilitative, and Health-Preserving Technologies, 3(1), 1–129.

[2] Hendriks, N., Huybrechts, L., Wilkinson, A., & Slegers, K. (2014). Challenges in Doing Participatory Design with People with Dementia. In Proceedings of the 13th Participatory Design Conference: Short Papers, Industry Cases, Workshop Descriptions, Doctoral Consortium Papers, and Keynote Abstracts - Volume 2 (pp. 33–36). New York, NY, USA: ACM.

Workshop:
IFIP WG 13.2 Workshop on User Experience and User-Centered Development Processes

User Experience Centered Engineering: A Process Model Inspired by Games Development

Regina Bernhaupt*, François Manciet* and Michael Pirker**
(*) IRIT
ICS Group
118, Route de Narbonne, 31062 Toulouse, France
{Regina.Bernhaupt, Francois.Manciet}@irit.fr
(**) ruwido
Köstendorferstr. 8, 5202 Neumarkt, Austria
{Michael.Pirker}@irit.fr

Abstract. Following a brief review of standard software engineering processes and examples of software developments in the games industry, a user experience (UX) centered process model is proposed highlighting a set of methods to be performed for the various phases (UX Concept, UX Design, Prototype, System Production, Post Production) and the iterative (and repetitive) stages (Analysis, Design, Implementation, Evaluation).

1 Introduction

User Experience (UX) is defined as "a person's perception and the responses resulting from the use or anticipated use of a product, system, or service." [7] . In the last 20 years, UX was investigated using a variety of definition ranging from approaches that were rather holistic to definitions that were related to one or few concepts. McCarthy et al. [4] argue that UX is a holistic term, as the sum of a set of factors or concepts can be more than just the individual parts. On a more practical viewpoint, it is clear that UX is made up of a set of factors or concepts that we can develop for, and that are more tangible than investigating such a holistic experience (Mc Carthy, personal communication). A set of (well chosen) factors seems to be a good starting point, to engineering for a better UX. User experience is described as dynamic, time dependent [5] and beyond the instrumental [6] . From an HCI perspective, the overall goal of UX is to understand the role of affect as an antecedent, a consequence and a mediator of technology. The concept of UX focuses rather on positive emotions and emotional outcomes such as joy, fun and pride [10] .

From a software engineering viewpoint, a factor or concept oriented definition of UX is a good starting point. Individual factors or a set of factors

can be (easier) engineered for, when developing for a non-specified "experience".

When developing software that focuses on UX, the user and the experience the user has when interacting with the system are of key importance. In the following we propose an adapted UX-centered design and development process with a set of related methods that can help to engineer for a set of UX dimensions or factors.

2 User Centered Process Model

When engineering for UX there are some differences in the design and development cycle compared to standard software engineering models. To be able to involve the user at all stages of the development process, a user centered development process is a good starting point. Figure 1 shows such a standard user centered development process (adapted from [1] , p. 74) enabling iterations.

Project preparation phase / Concept phase

	Design →	**Implement** →	**Evaluate**
Concept design	Initial design concept	Design mock-ups	Interviews / reaction to mock-ups
Functionality/UI design	Design of system functionality / look and feel	Low fidelity prototypes	Usability study
Prototype implementation	Design of critical mass of system	Implementation of partial working system	Heuristic evaluation
System enhancement and evolution	Design of monitoring and feedback system	Implementation of monitoring and feedback system	Interviews and data collection

Figure 1. Iterative user centered design and development cycle, adapted from [1]

The process from Figure 1 is still useful today, but has been extended in the fields of Human Computer Interaction (HCI) and Software Engineering (SE) to include a variety of phases and methods depending on the type of interactive system being built. For example, for the development of large systems it is important that project preparation phase and concept design today

typically include the gathering of requirements (also to support requirements engineering). In the waterfall model [3] , requirements engineering is presented as an early phase of the development process. Later software development methods, including the Rational Unified Process (RUP), Extreme Programming (XP) and Scrum assume that requirements engineering continues through the lifetime of a system.

The waterfall model refers to the following phases for SE development: system feasibility, software plans and requirements, product design, detailed design, code, integration, implementation and operations and maintenance [1] (p. 282). What is clearly missing is the ability to iterate and the ability to evaluate.

When focussing on UX the concept and the early phases of design and development can be structured differently. We discuss in the following a standard iterative design and development cycle consisting of (2.1) early design and development phase, (2.2) prototype phase, (2.3) implementation phase several phases and methods that are special when focusing on UX.

2.1 Early Design and Development Phases

Early Design and Development Phases can be referred to as the project preparation phase or concept phase (see Figure 1). When designing for UX, the number and type of participants and competencies can be challenging, as UX is a process that is typically not only targeting at the software itself, but at other processes that can influence the experience of the user. The experience we want to design for is not similar to the user interface (UI) design, but reaches beyond the UI. For the development of new products, this can include industrial design to support the form factors and shapes of the product (e.g. the size and form of the screen on the mobile phone, remote control or shape of the intelligent alarm clock), it can involve material engineering to develop new materials with special abilities (e.g. flexible screens [11] , materials that store energy to support wearables [8] ...), it typically includes marketing and public relations (e.g. to develop a brand identity the users can identify with), and so on.

When working with such an interdisciplinary team to identify the type of UX the product shall deliver, simple textual descriptions for requirements are not enough. The communication of UX visions and ideas can be done for example verbally using a method called "Elevator Pitch" where you describe the idea within the time it takes the elevator to reach the floor, or to use only one word to focus on the topic. Other possibilities are to use metaphors,

experiences everyone can relate to (e.g. [9] mentions "to use light like on a dentist chair"), or to use one question.

Non-verbal descriptions include HCI-related methods that can range from Mood Boards [9] , to pictures and concept art (typically rough sketches). They can be complemented by videos showing sample animations.

Specificities for Games and Entertainment: In games development the concept phase "begins when an idea for a game is envisioned - and it ends when a decision is made to begin planning the project" (see [12] , p. 334). This phase is dedicated to the initial game idea and is devoted to producing a first concept document describing the game. The development team in this phase is typically small (e.g. consisting of designer, programmer, artist and producer). In the concept phase the game play experience you want to provide and the target market are identified. The concept phase is followed by the pre-production phase, where a game proposal and a planning are developed as well as additional documentation is produced (art style guide, production plan). The phase ends with the production of the game design document and the technical design document.

2.2 Prototype Phase

Goal of the prototype phase is to create something tangible (see *Concept design* and *Functionality / UI design* in Figure 1). Prototypes can range from low-fidelity prototypes showing main ideas on paper or in a power-point presentation to high-fidelity prototypes that already allow the evaluation of UX dimensions like aesthetics, emotion, identification, stimulation or social connectedness [2] .

Specificities for Games and Entertainment: Goal of this phase is a first working piece of software allowing to demonstrate key characteristics of the game and enabling to understand basic concepts related to the general UX of the game (“Is the game fun to play?”). This requires high-fidelity prototypes that are typically showing the technological aspects that will be new for the system as well as the game play experience. They are also used to show the game idea to a game developer study to acquire funding.

2.3 Implementation Phase

For the implementation of the interactive system there is a variety of SE methods and approaches available that help support a development that is iterative and ready for change (based on evaluation results, see *Prototype implementation* in Figure 1). Methods include SCRUM, agile development and others.

Specificities for Games and Entertainment: The implementation phase for games is typically referred to as production phase and has some special phases:

- Alpha-Phase: This is the phase when a game is playable from start to finish, allowing different evaluation methods to be applied to better understand aspects like fun, playability and UX.
- Beta-Phase: Main goal during this phase is normally to fix bugs. In terms of UX, in this phase lots of fine-tuning is necessary to improve the overall UX. The beta-phase includes steps like certification or submission (i.e. the hardware-manufacturer of the proprietary platform will test the game).
- Gold: In this phase the game is sent to be manufactured.
- Localization: An important phase for games that will be delivered to different markets (countries) is the localization phase. In this phase game-play can be adjusted to suit the tastes of the market, to allow for language translation and modifications due to local regulatory authorities.

2.4 After Release

After a system was released on the market, the usage of the system can be monitored (see *System enhancement and evolution* in Figure 1). On the basis of the results from such a monitoring-based evaluation, the system properties and functions can be adapted.

Specificities for Games and Entertainment: This phase is called post production and is becoming more and more important, as there is lot of activity to balance the game play, to improve and extend the number of game options or levels and so on (see [12] , chapter 5 for more).

3 A User Experience Centered Process Model

When developing for UX, it is important to follow an iterative design and development process. Following the four major steps of analysis, design, implementation and evaluation it is possible to react quickly to incorporate necessary changes to improve the UX.

To analyse UX there is a variety of methods. Standard social science methods like focus groups, interviews and observation can be used at early stages to discover what type of experience we design for and how the context the product is used in is influencing experiences. For the design methods and

approaches like mood boards, videos and photos as well as high-fidelity prototypes help to design for the experience that is envisaged.

For the implementation there has recently been a lot of development (especially in games) enabling to model different users to support individual play styles [13] , to analyse the history of interaction with the game to vary the difficulty of the game as well as interaction mechanisms that help to cheat if people are playing together.

For the evaluation of UX there are three types of methods available: expert-oriented methods, user-oriented methods and automatic methods [2] .

As Figure indicates for a UX-centered design and development, we propose to follow the four iterative steps including *analysis*, *design*, *implementation* and *evaluation*. The process model consists of five phases: (1) the *UX concept* phase with the main activity of analysing the usage context, users and how this will affect the UX dimensions you want to support with the system, (2) the *UX design* phase that will (contrary to standard user centered design and development) consist of several iterations including low, but also high-fidelity prototypes and sometimes even partly functional systems to study UX. (3) The *Prototype* phase includes the first tuning on UX dimension selection, and might lead to shifting to a different experience. Possible evaluation methods at that phase will be expert evaluations (and insight from marketing) checking if the initial premises on how to support the UX dimensions were successfully applied. (4) The *Production* phase will typically be repeated - for games, production can be up to 2 years with hundreds of programmers, so tradition SE approaches are combined with agile methods as well as approaches like regular builds by the end of the week or first playable version, alpha and beta release as milestones. (5) Today, the *Post-production & evolution* phase is one of the most important phase for interactive systems, as they are continuously changing and developing due to new features, extensions or changed expectations of the users. Websites like Facebook or Amazon are changing on the fly, continuously enhancing the UX.

Project preparation phase
Identification & description of UX dimensions

	Analyze →	**Design** →	**Implement** →	**Evaluate**
UX concept ↓	Contextual inquiry	Sketches	Mood Boards, elevator speech...	Evaluate question
UX design ↓	UX dimensions	Low fidelity	Paper prototypes / mock-ups	Interviews
		High fidelity	Vertical prototype	UX study

Prototype ↓	UX dimensions tuning	Design of critical mass of system	Implementation of partial working system	Heuristic evaluation
Production ↓ (can be repeated)	Fine tuning	Development of deliverable system	Implementation & installation	Usability tests
Post-production & evolution ↓	UX enhancement	Monitoring & feedback system	Implementation of mon. & feedback system	Interviews and data collection

Figure 2. Proposed Phases and Methods for UX-oriented design and development of interactive systems.

4 Conclusion

Including UX as central driver for software development is a difficult endeavour. This position paper argues for a five phases UX-centered development process that integrates the typical iterative design and development phases of analysis, design, implementation and evaluation. Main goal of this work is to lay a foundation for discussion within the workshop to identify gaps of the process model, to enable gathering of methods that are applicable and to vividly discuss how to support UX with engineering methods like user models, adaptive systems, personalisation and individualisation to support that user experiences are changing over time, based on the context the system is used in, and are individual to each user.

References

[1] Baecker, R. M., Grudin, J., Buxton, W., Greenberg, S. Readings in Human-Computer Interaction: Toward the Year 2000 (Interactive Technologies). Morgan Kaufkman, 1995.

[2] Bernhaupt, R. ed. Game User Experience Evaluation. Springer Verlag, Heidelberg, 2015

[3] Boehm, B., W., A Spiral Model of Software Development and Enhancement. Computer 21, 5 (May 1988), 61-72. DOI=10.1109/2.59 http://dx.doi.org/10.1109/2.59

[4] McCarthy J, Wright P. Technology as Experience. MIT Press, Cambridge, MA, 2004.

[5] Karapanos E, Zimmerman J, Forlizzi J and Martens, J-B. Measuring the dynamics of remembered experience over time. Interacting with Computers 22(5), 328-335, 2008.

[6] Hassenzahl M The Thing and I: Understanding the Relationship between User and Product. In: Blythe MA, MOnk AF, Overbeeke K, Wright PC, (eds) Funology: From Usability to Enjoyment, Kluwer Academic Publishers, Netherlands, 2003.
[7] ISO 9241-210,
[8] Jung, S., Lee, J., Hyeon, T., Lee, M. and Kim, D.-H. Fabric-Based Integrated Energy Devices for Wearable Activity Monitors. Adv. Mater., 26: 6329–6334, 2014.
[9] Hagen, U., Designing for Player Experience: How Professional Game Developers Communicate Design Visions. DiGRA Nordic Proc., 2010. http://www.digra.org/wp-content/uploads/digital-library/10343.03567.pdf
[10] Hassenzahl M and Tractinsky N (2006) User Experience - a research agenda. In: Behavior & Information Technology, 25(2), 91--97.
[11] Mone, G. The future is flexible displays. Commun. ACM 56, 6 (June 2013), 16-17. 2013.
[12] Novak, J. Games Development Essentials. Second Edition. Delmar Cengage Learning, 2009.
[13] Yannakakis, G.N., Togelius, J., Experience-Driven Procedural Content Generation, In: Affective Computing, IEEE Transactions on, vol.2, no.3, pp.147,161, July-Sept. 2011 doi: 10.1109/T-AFFC.2011.6

Mining Logs to Support HCI (Re)Design

Felipe Cordeiro de Paula, Simone D.J. Barbosa
Departamento de Informática, PUC-Rio
Rua Marques de Sao Vicente, 225
Gavea, Rio de Janeiro, RJ, Brazil, 22451-900
{fpaula, simone}@inf.puc-rio.br

Abstract. This position paper proposes to investigate the use of logs as a resource to (re)design interactive systems. By understanding what information should be captured in application logs and using mining techniques to identify the users' behavior patterns, we hope to extract valuable information to inform HCI (re)design decisions.

1 Introduction and Motivation

There are many languages, technologies, and services available for designers and developers to create interactive systems. Choosing from among these requires knowledge to find a suitable, contextualized solution and to avoid producing inconsistent artifacts or a system with poor interaction design.

A good systems designer should care about the interaction between users and computers, based on the principles of efficiency, effectiveness and satisfaction [3]. In this context, the area of Human-Computer Interaction (HCI) investigates the "project (design), evaluation and implementation of interactive computing systems for human use, and the phenomena related to this use" [1].

Having in mind these three principles, the motivation of this proposal stems from the potential to analyze the interaction process recorded in log files, using data mining techniques, in order to identify usage patterns, inconsistencies or requirements not adequately met by the system, to promote HCI (re)design.

In general, our challenge involves addressing the following questions: (i) What information must exist in log files structures so that designers can apply data mining techniques to discover gaps and opportunities for building systems that adhere to sound HCI principles and guidelines? (ii) How to analyze that info to investigate the user behavior, including user choices of interaction and their frequencies, in order to make transparent the habits, vices, and even system options that are set aside and left unused?

2 Justification

A great variety and amount of events can be observed on a system through its log files, but most systems record only critical events and access control. In these traditional logs, we already find information about the operation of programs, data errors, or routine checks. Taking as an example the Microsoft Internet Information Server — in W3C format —, we see it provides the following log info: date, time, client IP address, method, source URI, URI query, the number of bytes sent, the HTTP status and the protocol version. All these data are important, but they do not directly assist us when thinking about interface redesign.

We would therefore like to explore how these logs could record different aspects of user behavior, going beyond security or critical operations. Log files should allow for continuous monitoring, allowing designers to reflect on the navigation choices, bypasses and habits, assisting them to identify what are the events that occur more frequently, the underutilized steps and resources [4][5][6]. Enhancing the use of log files beyond current practices is at the heart of our proposal.

Researchers on game design and development have successfully defined and used analytics approaches to better understand gamers' behavior and adapt game applications, both at design and run-time [1].

This proposal thus tackles the use of log files from an HCI perspective, where a new vision for log creation and interpretation is needed, allowing us to improve software quality by examining various kinds of events (e.g., events that trigger the execution of a task and detailed steps that comprise a task or process.

3 Preliminary Solution Approach

We propose to build a log file template to record additional info to support HCI (re)design activities and help improve the human-computer interaction, using data mining techniques. To achieve this, we first need to define desirable characteristics of a log file, with the ability to adapt to different domains, and then follow these steps:

- Register the different interaction paths followed by users to perform their activities;
- Analyze the user's behavior and use of the system;
- Develop a prototype or modify an existing system to use the proposed log format;
- Analyze the results obtained from the use of this prototype;

- Compare the design processes followed by those who used the prototype (with the enhanced log) with a control group that did not use it.

References

[1] El-Nasr, M.G., Drachen, A. Canossa, A. Game Analytics: Maximizing the Value of Player Data. Springer, 2013.

[2] Hewett, Thomas T. et al. ACM SIGCHI curricula for human-computer interaction. ACM, 1992.

[3] ISO 9241-210 - Available at: https://www.iso.org/obp/ui/#iso:std:iso:9241:-210:ed-1:v1:en – Accessed on 8 June 2015

[4] Van Der Aalst, W. M. P., Weijters, A. J. M. M., (2004), "Process mining: a research agenda", Computers in Industry, v. 53, n. 3 (Apr.), p. 231-244.

[5] Van Der Aalst, W. M. P., Weijters, A. J. M. M., Maruster, L., (2004), "Workflow Mining: Discovering process models from event logs", IEEE Transactions on Knowledge and Data Engineering, v. 16, n. 9, p. 1128-1142.

[6] Van Der Aalst, W. M. P.; Günth, C. W. (2007) “Finding structure in unstructured processes: The case for process mining”. In: Application of Concurrency to System Design. ACSD 2007. Seventh International Conference on. IEEE, 2007. p. 3-12.

Integrating Human-Centered and Model-Driven Methods in Agile UI Development

Holger Fischer, Enes Yigitbas, Stefan Sauer
s-lab – Software Quality Lab
University of Paderborn
Zukunftsmeile 1, 33102 Paderborn, Germany
{hfischer, eyigitbas, sauer}@s-lab.upb.de

Abstract. The development of interactive systems is a challenging task and requires structured methods and processes for developing high quality user interfaces. Existing approaches for developing interactive systems are mainly based on the model-driven user interface development (MDUID) and human-centered design (HCD) paradigms. Although these approaches support the efficient development of interactive systems considering human factors, they lack in flexibility to cover dynamic characteristics in the development process. To overcome this deficit we propose an agile UI development process that combines strengths of MDUID and HCD and at the same time enables an incremental development process having a continuously runnable part of the software front-end ready for reviews with the end user.

1 Motivation

Today, the usability of an interactive system has been recognized as an important quality aspect within software development industry. The development of the interaction between humans and interactive systems is a challenging task and requires structured methods and processes for developing and maintaining high quality user interfaces.

Model-driven development (MDD) [1] is a current development paradigm that addresses the creation of domain-specific and formalized models with the aim of generating a code base. Abstraction is a key factor of MDD to describe domains regardless of the target programming language or platform. Hence, problem descriptions can be specified with more precision and free from redundancies. Furthermore, the traceability of changes and the reusability of models are supported using specific tool chains for transformation and generation. However, MDD only focuses on the users in the initial analysis. There is no continuous participation of users during the development stages. A formative or summative evaluation of the specified models or the user interface is not in focus.

Human-centered design (HCD) [2] addresses such continuous user participation during design and development. User requirements and design solutions are refined over the time using an iterative design as well as an early focus on the users with their needs, goals and tasks. Integrating users during the whole development enables a continuous feedback and validation instead of interpretation and assumptions. Hence, obtained insights of user requirements analysis and usability tests are usually specified in a narrative way using scenarios, storyboards or reports instead of formalized models. Furthermore, HCD focuses on the development of design solutions, but not on the incremental development of UI components and the underlying system.

Agile development (e.g. Scrum [3]) enables an incremental development of software having a continuously runnable part of the software ready for reviews with the customer. Therefore, transparency and inspection are two key factors to create a common understanding within a project. However, agile development is not targeted on modeling user requirements and human behavior in a comprehensive way. Having inspections with the customer does not fulfill the expectations of having a broad analysis and evaluation with multiple users who are currently working with a system. Additionally, a systematic and documented way of decisions concerning the interaction or the UI as well as a systematic treatment of user feedback is missing.

The existing gap between human-computer interaction (HCI) and software engineering (SE) is still a challenge as well as a chance to define a systematic method for software development. The challenge addressed in this paper is to create a model-driven method for agile UI development that combines the advantages of all three paradigms mentioned above: Systematic and sustainable model-driven development, early and continuous user participation as well as incremental deployment. Thus, the open issues in every single paradigm will be fixed within their combination.

The paper is structured as following: First, we introduce some related work on the topics of model-driven HCI development and agile model-driven development. Thereafter, we present the concept of our agile model-driven UI approach. Finally, we conclude our paper and give an outlook for future work.

2 Related Work

Focusing on the topic of agile UI development integrating human-centered and model-driven methods, multiple aspects have to be taken into account. In the following we will briefly sum up existing approaches on model-driven UI development and human-centered design as well as agility in model-driven development approaches.

Model-driven User Interface Development (MDUID) brings together two subareas of software development, which are MDD and user interface development (UID). The core idea behind MDUID is to automatize the development process of UI development by making the models the primary artifact in the development process rather than application code. An MDUID process usually involves multiple UI models on different levels of abstractions that are stepwise transformed to the final user interfaces by model transformations. The CAMELEON Reference Framework (CRF) [4] provides a unified reference framework for MDUID differentiating between the abstraction levels task & concept, abstract user interface (AUI), concrete user interface (CUI) and final user interface (FUI).

There are various state-of-the-art modeling languages for covering the different abstraction levels of the CRF. For example MARIA XML (Model-based lAnguage foR Interactive Applications) [5] and IFML (Interaction Flow Modeling Language) [6] provide both an AUI modeling language and a tool-support to create and edit AUI models. Based on these AUI models further transformations can be performed to transform them into platform-specific CUI models, which eventually are needed for generating the final user interfaces (FUI). The described MDUID approaches enable the specification and also support the generation of UIs, but they do not target aspects of HCD. Therefore recommended activities of the HCD process under the terms of ISO 9241-210 are not sufficiently integrated in the MDUID approach. To overcome this deficit a model-based HCI development process was proposed by Petrasch [7]. Here the author proposes a HCD process that is based on formalized models to cover the different aspects of the HCD process like describing the context of use or specifying user requirements. Although this approach tries to align model-based UI development methods with HCD aspects, it is not focusing on the topic of agile UI development.

In order to combine MDUID and HCD in an agile manner it is important to analyze existing approaches on agile model-driven development. In this context our proposed method (see section 3) is based on the agile model-driven development (AMDD) approach by Ambler [8]. As the name implies, AMDD is the agile version of MDD. MDD is an approach to software development where extensive models are created before source code is written. The difference with AMDD is that instead of creating extensive models before writing source code you instead create agile models, "which are just barely good enough that drive your overall development efforts" [8]. The AMDD lifecycle starts with an envisioning phase where initial requirements are specified and an initial architecture envisioning is sketched. After this first iteration the actual development iterations start. Development iterations

consist of a model storming, modeling iteration and test-driven development phase. An iteration phase can be guided by optional reviews of the developed increment. The development iterations are repeated until a software product results that conforms to the user's requirements.

3 Integrated Development Method

Taking model-based HCI development and agile model-driven development into account, we propose an integrated approach (Figure 1) with a specific model flow and evaluation feedback (Figure 2) to combine

- a systematic and sustainable model-driven development,
- early and continuous user participation and
- incremental deployment.

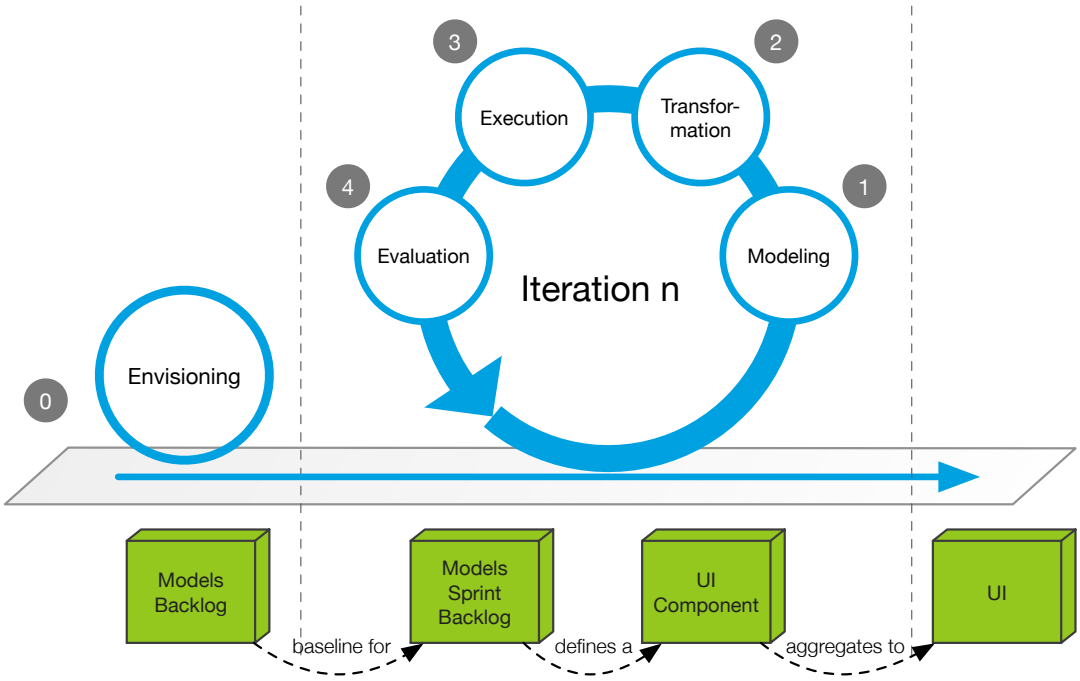


Figure 1. Agile model-driven UI development lifecycle

This approach consists of four stages and an additional envisioning up-front stage, which are described as follows:

0) Envisioning

In the envisioning up-front stage, a business analyst analyzes the context of use considering the users with their characteristics, needs, intentions, tasks, physical and social environment as well as the business goals of the organization. Then, all gathered information are described in specific models,

e.g. user model, task model, domain model, platform model, environment model. These models build up the models backlog corresponding a product backlog in agile development. This stage serves as a baseline for the development and will be extended in the modeling step during iterations.

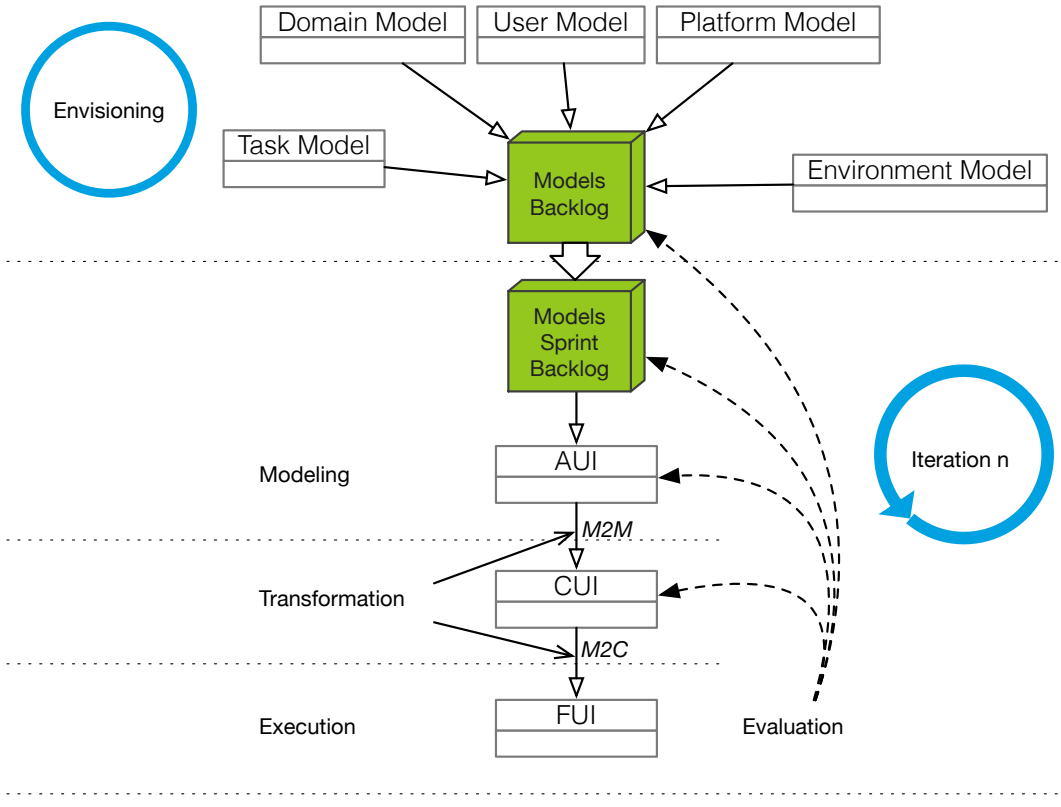


Figure 2. Model flow and evaluation feedback

1) Modeling

Each iteration (or sprint) starts with a modeling stage. Similar to a sprint backlog, the relevant models out of the models backlog will be expanded, specified as well as refined in more detail as an AUI model. This is conducted by using modeling languages like the interaction flow modeling language (IFML) [6] or MARIA XML [5] and may be enriched with formalized UI patterns like GUI patterns (e.g. wizard or auto-completion pattern) [9].

2) Transformation

Using model-to-model (M2M) transformations, the AUI model is than transformed in a CUI model specified for the target platform. Based on the CUI model the initial code for the FUI can be generated with a model-to-code (M2C) transformation.

3) Execution

Afterwards, a generated code base will be implemented with the functionality needed and which will be able to enrich with business logic. The previously implemented software component will be deployed and executed.

4) Evaluation

During the execution of a software component the behavior of the different test users will be logged in a pseudonymized way using appropriate logging mechanisms. The logged data include e.g. interaction flow, time needed to complete a task, used interaction objects or the frequency of performed tasks. These information will be analyzed e.g. to identify shortcuts in navigation paths (elephant paths), to restructure the navigation based on often performed tasks or to identify interaction barriers out of time deviations. In addition, usability tests may be carried out to gather more qualitative feedback. The model backlog will be adapted with the gathered insights of the evaluation in order to start the next iteration. According to the kind of feedback the results will be iterated on the appropriate abstraction level of the underlying models. If there are minor changes concerning the representation on the platform level the CUI models will be enhanced. Otherwise, conceptual changes have to be revised in AUI models or even further in the task and concept models.

4 Conclusion and Outlook

In this paper, we presented a concept of an agile user interface development approach enriched with human-centered and model-driven methods. Thus, we try to combine the advantages of all three paradigms in one approach: An increased usability through HCD methods and user participation; an improved sustainability using model-driven methods; continuously runnable software parts for a better communication within the development team and towards customers and users.

We believe that this combination is promising to software developing enterprises rapidly being able to create runnable and usable software releases. In this way, we focus on aspects to change software engineering processes to support usability as well as to synchronize HCD with software development activities based on fast sprints.

In our future work we will further expand the concepts in more detail. In doing so, we are thinking about logging mechanisms to study the users' behavior. Gathered data could then be abstracted within formalized models. These models could be compared with the already developed models in the backlog in order to emphasize and improve the differences between these models.

In addition, we will evaluate the concept within small software projects as well as within our teaching activities.

References

[1] Stahl, T. and Voelter, M. Model-Driven Software Development: Technology, Engineering, Management. John Wiley & Sons Ltd., Chichester, UK, 2006. ISBN 978-0470025700

[2] ISO 9241-210:2010 Ergonomics of human-system interaction – Part 210: Human-centered design for interactive systems, 2010.

[3] Schwaber, K. and Sutherland, J. The Scrum Guide, 2013. http://www.scrumguides.org/ Last visit: 05/10/15.

[4] Calvary, G., Coutaz, J., Thevenin, D., Limbourg, Q., Bouillon, L. and Vanderdonckt, J. A Unifying Reference Framework for Multi-target User Interfaces. In Interacting with Computers, Vol. 15. Elsevier, 2003. pp. 289-308.

[5] Paternò, F., Santoro, C. and Spano, L.D. MARIA: A universal, declarative, multiple abstraction-level language for service-oriented applications in ubiquitous environments. In ACM Transactions on Computer-Human Interaction, Vol. 16 (4), Article 19. ACM, New York, 2009.

[6] Interaction Flow Modeling Language (IFML), Specification 1.0. OMG, 2015. http://www.omg.org/spec/IFML/ (Retrieved June 11, 2015)

[7] Petrasch, R. Model based User Interface Design: Model Driven Architecture and HCI Patterns. In Softwaretechnik-Trends 27 (3), 2007.

[8] Ambler, S.W. Agile Model Driven Development (AMDD): The Key to Scaling Agile Software Development. http://www.agilemodeling.com/essays/amdd.htm (Retrieved June 11, 2015)

[9] Yigitbas, E., Mohrmann, B. and Sauer, S. Model-driven UI Development integrating HCI Patterns. In Proceedings of the 7th ACM SIGCHI Symposium on Engineering Interactive Computing Systems (EICS), International Workshop on Large-scale and model-based Interactive Systems: Approaches and Challenges. 2015 (to be published)

Managing the Agile Process of Human-Centred Design and Software Development

Peter Forbrig[1] and Michael Herczeg[2]
[1]University of Rostock, Institute of Computer Science
Albert-Einstein-Str. 22, 18051 Rostock
peter.forbrig@uni-rostock.de
[2]University of Lübeck, Institute of Multimedial and Interactive Systems
Ratzeburger Allee 160, 23562 Lübeck
herczeg@imis.uni-luebeck.de

Abstract. The paper provides and discusses a life cycle model for agile software development methods like SCRUM. The process model integrates the idea of Human-Centred Design (HCD) in the agile approach. Based on the discussion of existing life cycle models and known process patterns, a process model is discussed that combines the advantages of HCD and SCRUM. The SCRUM process model was used in the discussion to be as concrete as possible. However, the suggested approach is applicable to other agile process models as well.

1 Introduction

It is generally accepted that software should be developed in an iterative way because otherwise it is difficult to capture all requirements in a sufficient way. Even that the classical waterfall model has some advantages for the definition of milestones it is agreed, that the spiral model of Boehm and the agile software development together with formative evaluation methods provide better software quality in the sense of usability and user experience.

2 Related Work

Agile software development has become popular since the late 1990s. It became popular because many projects failed because it took too much time from finalised requirements specification to first tests of the developed system. The approach was consolidated by a manifesto at the beginning of ten 2000s [1] that characterises the agile idea by twelve main principles that have to be followed.

Figure 1 gives additionally an overview of influences to agile software development. It presents approaches from planning, analysis, design, build, test deploy and review that found their way into agile development methods.

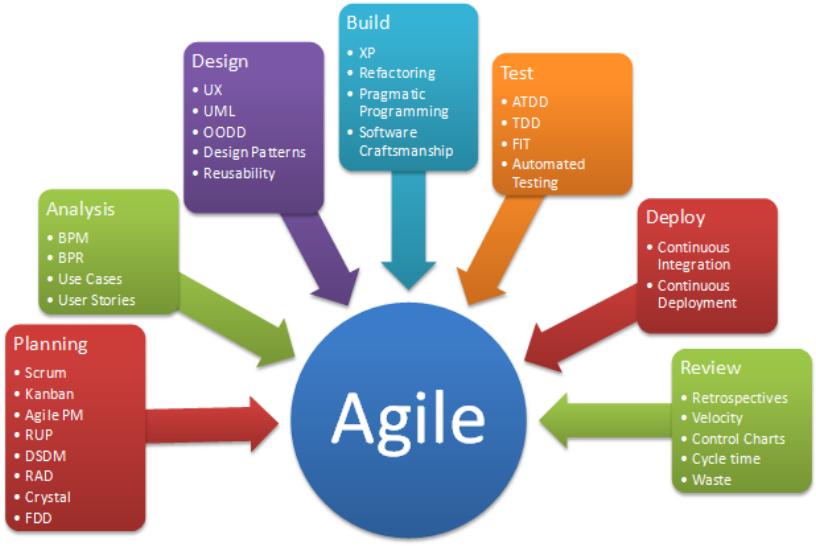


Figure 1. Agile influences (from http://blog.kiandra.com.au/wp-content/uploads/2013/08/Agile-21.png)

Currently, one of the most popular agile software development approaches is SCRUM, which will be revisited in the following paragraph.

2.1 Agile Software Development - SCRUM

To make things more concrete we focus on one specific agile development method. One of the most popular methods is SCRUM [9]. Therefore we will focus on the related process models and the corresponding terminology.

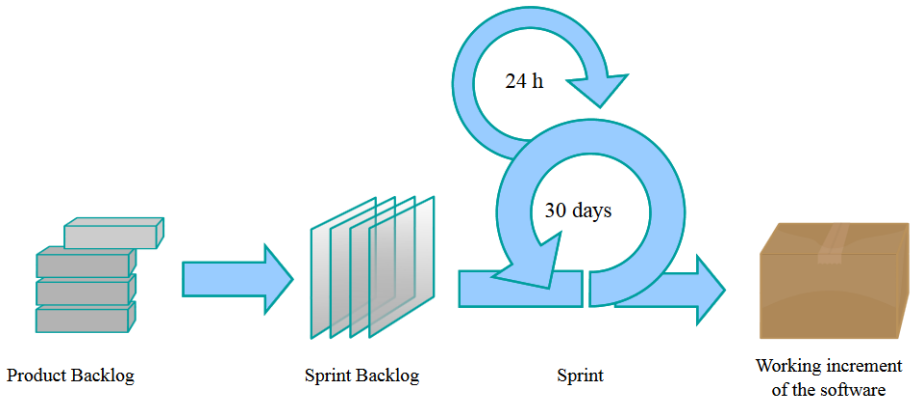


Figure 2. The scrum process (from http://en.wikipedia.org/wiki/Scrum_software_development)

Figure 2 presents the main process model for SCRUM. Based on the specified product backlog collecting all requirements for the product, a sprint backlog is selected that contains all requirements that have to be implemented during the forthcoming sprint. The result is an increment of the necessary software.

2.2 Human-Centred Design (HCD)

In the same way that SCRUM is popular for software engineering experts, HCD is popular for usability and user experience experts. One of the main reasons for its success is that context of use, the requirements of the users and the evaluation of design solutions play an important role. User requirements are more important than technical features that software engineers might derive. Users get what they really want.

The HCD process has been standardised by ISO 9241-210. Figure 3 gives a visual overview of the corresponding process model.

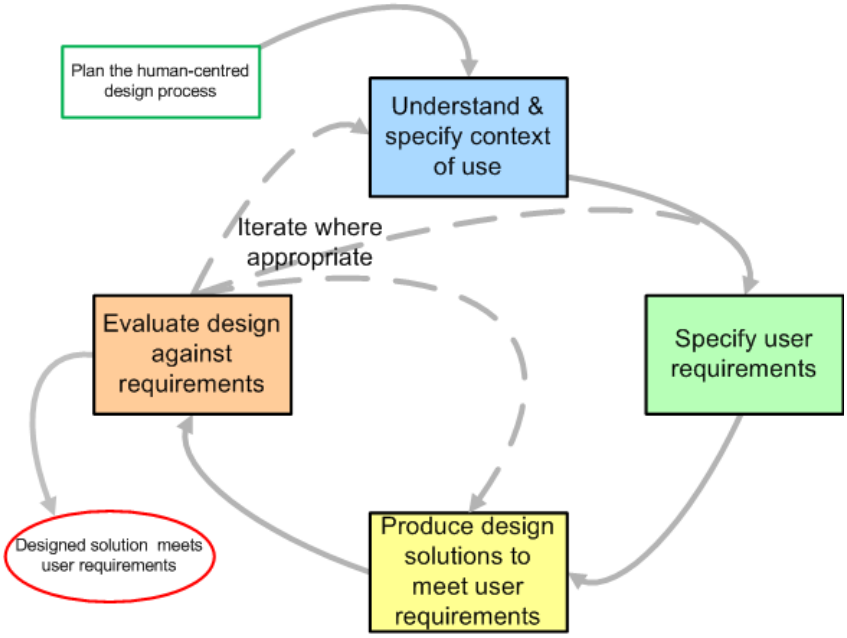


Figure 3. The design process from ISO 9241-210– Human-centred design process (from https://thestandardinteractiondesignprocess.wordpress.com/)

At the University of Lübeck a Usability Engineering Repository (UsER) was developed [7]. It is a framework, platform and repository for an integrated and modular development of human-machine systems with the special scope of usability engineering.

UsER supports the analysis, design and evaluation of interactive systems through interrelated analysis and design entities creating a meshed specification of the system to be developed.

The repository provides a broad variety of standard or specific methods of usability engineering. It enables the process and motivates the teams for user-centered development and user interface design thinking.

UsER is in the stage of an advanced prototype that has already been applied within the industrial development of ERP systems (business applications) as well as supervisory control systems (safety critical systems). It is a modular system that allows selecting methods for analysis, design or evaluation as needed for a specific development process.

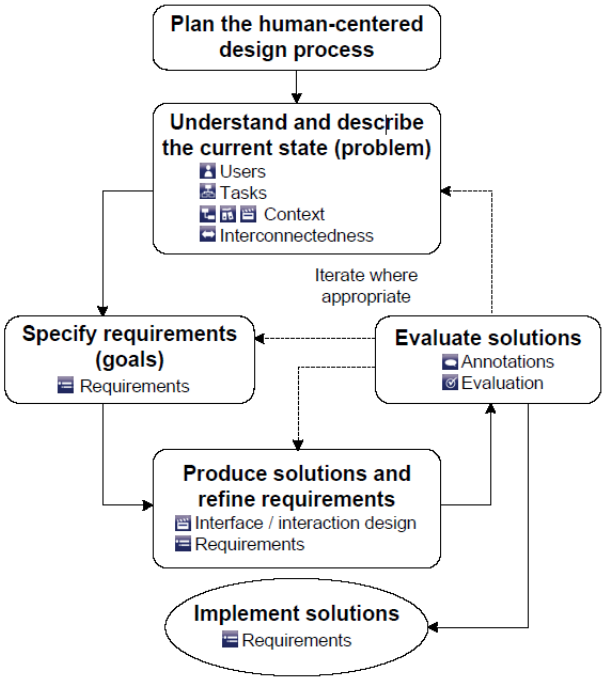


Figure 4. Process Model for UsER (from [4])

Figure 4 presents the basic process model of UsER. It extends the HCD process model by methods for modelling certain usability aspects. The icons represent certain modules that allow the specification of related models. To understand the current problem one is able to perform a user analysis and task analysis and to store the corresponding models. The context can be modelled by an organizational analysis, artifacts and/or scenarios. It is also possible to model the interconnection of different models. It can be easily seen that UsER is an appropriate tool support for HCD since it provides the essential and established methods of usability engineering like user, task and context modelling within an integrated environment.

An integrated model for HCD and agile software development was developed in [8] and will be presented in the following paragraph.

2.3 Approaches for Combining HCD and SCRUM

It was already mentioned that a combined approach for HCD and SCRUM would be appreciated by a lot of software developers because such a unified approach would give the chance to combine the advantages of both approaches. Paelke et al. [6] published the following process model and called it "Agile UCD-Process". However, it is more a UCD-Process with agile software development. Nevertheless, it is a very interesting starting point.

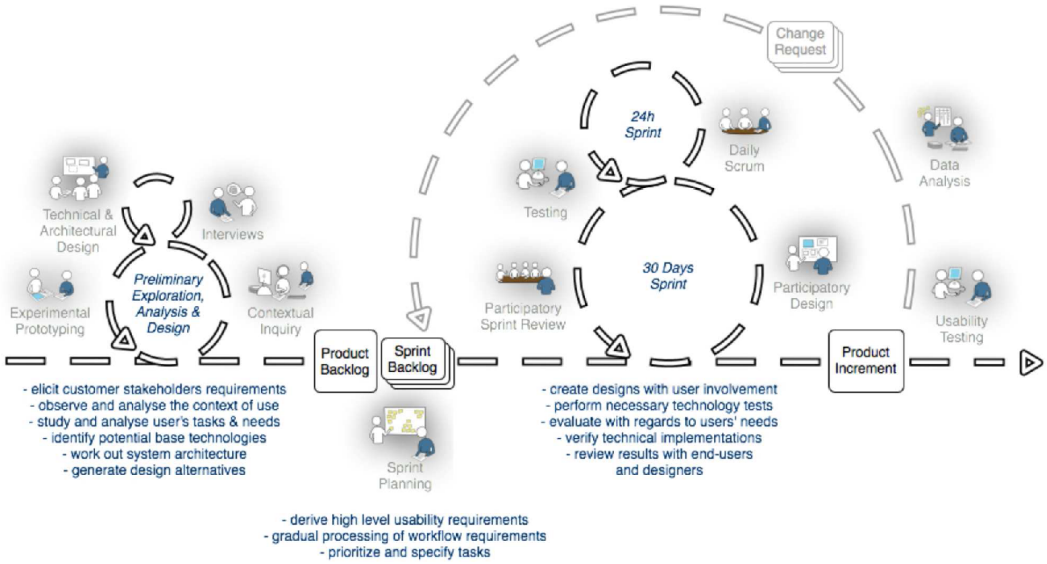


Figure 5. Agile User-Centred Design Process (Paelke et al. [5])

From our point of view it is interesting to start with UCD activities. However, the requirements elicitation is only loosely coupled with the software development process. A stronger coupling was suggested by Paul et al. [9] and is presented by the following Figure.

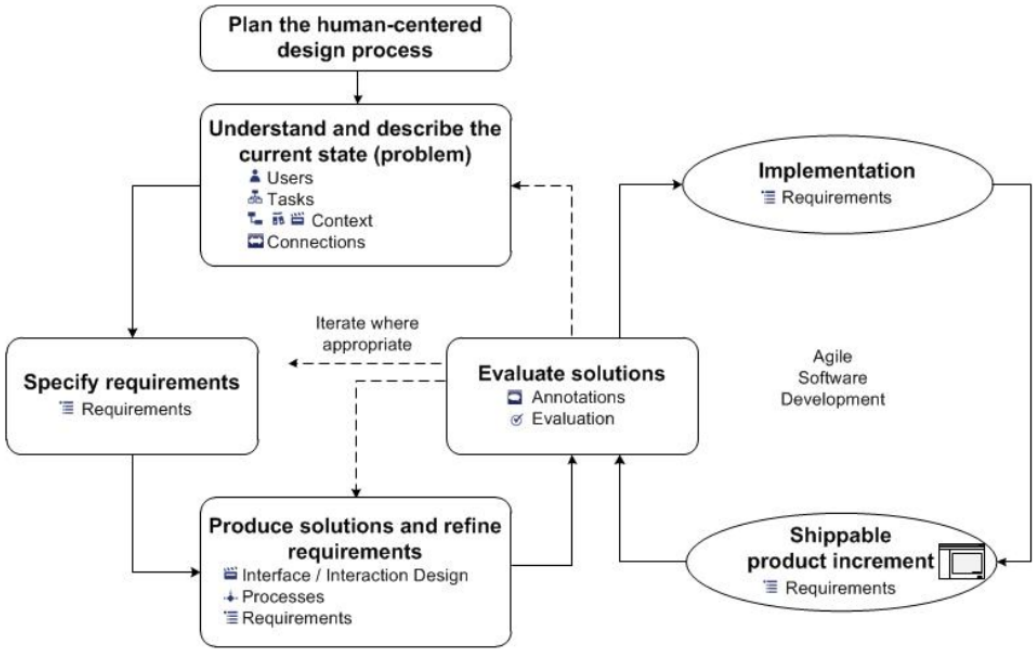


Figure 6. Extended User-Centred Design Process (Paul et al. [9])

According to this approach all created software influences the HCD process. U-SCRUM [11] is an approach with similar goals. It suggests to integrate usability aspects in SCRUM by having two product owners and to use the concept of personas for communication. However, the process is only roughly described by necessary artifacts.

Sinah et al [10] provide a practitioner perspective to the problem. They analysed that "*agile teams suffer from the additional challenge of lack of management support to UCD activities that occurred to participants*". They mention that a clear roadmap for combining both approaches is missing and that parallel tracks might help. We try to go one step forward to such a roadmap by presenting a process model in paragraph 2.5.

2.4 Process Patterns

Patterns were introduced to the computer science community by the "Gang of Four" [3]. Their collection of Design Patterns was a great success for reusing knowledge of solutions for reoccurring problems. In the meantime patterns for different domains like tasks, workflows or user interfaces were discovered and published. Sutherland et al. [13] published nine patterns related to SCRUM.
We will discuss with the pattern *"Yesterday's weather"* only one of them. This pattern says: "*In most cases, the number of Estimation Points completed in the last Sprint is the most reliable predictor of how many Estimation Points will be completed in the next Sprint*". The other patterns are of similar kind

and do not say a lot about the SCRUM process model. However, Bertholdo et al. [2] analysed projects and reports and provided patterns for agile software development that ask for consequences of the process model. The first pattern is called "Sprint Zero" and attacks the following problem: "Missing the big picture of the system in the beginning ....". It suggests the following solution:"*A short Sprint before the code implementation to define a broader view of the product, general goals, to roughly plan the next sprints and to define design principles*". This process model of Figure 5 can be seen as one variant of the consequences of this pattern.

Another pattern is called *"One Sprint Ahead"*. It is related to the following problem: "*Making the development and the UX team synchronized, that both can collaborate and provide input to the development workflow*". There is the following suggested solution:"*The UX team works at least one iteration ahead of the development team*".

The pattern *"Parallel Tracks"* provides the following solution: "*UX or usability team work in a parallel track with the development team in order to synchronise their activities*".

The suggested process model of Figure 6 is designed according to the solution of this pattern.

3 Updated Process Model for SCRUM

The result of our studies in projects studies and literature research is the following process model in Figure 7.

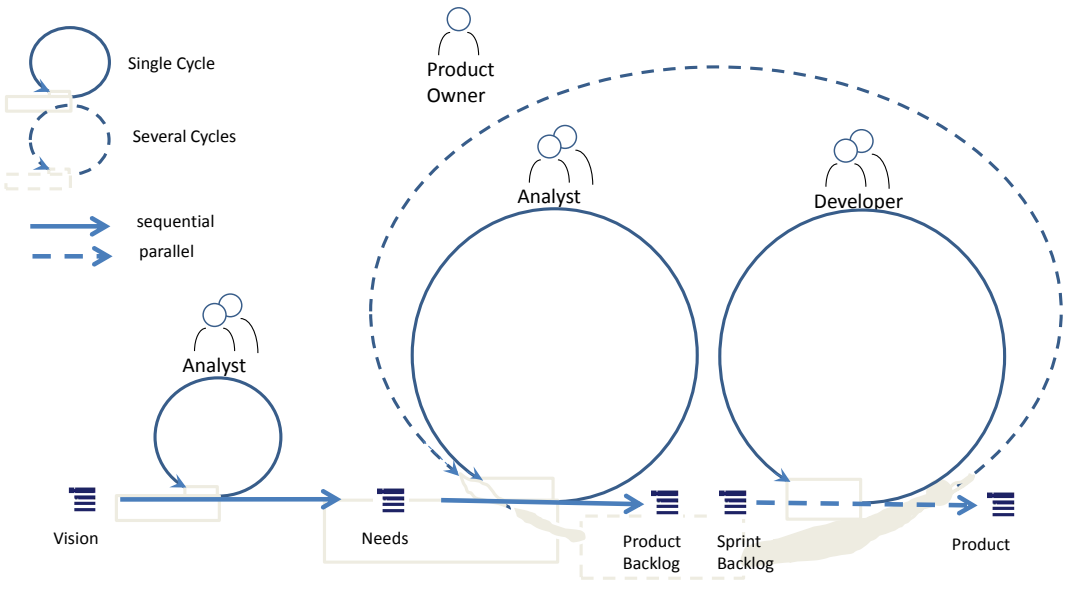


Figure 7. Suggested Human-Centred Design Process for SCRUM

The development process starts with a first analysis of the project context. Based on a project vision related needs are identified (pattern Sprint Zero). This activity is performed by people performing the analyst role. However, developers can perform this role as well and should be involved in the analysis process.

Based on the identified needs, requirements in the form of a product backlog are identified within the next sprint. Later the development sprint produces a product increment. This development Sprint is performed in parallel to the next cycle of identification of requirements in form of a product backlog increment. The dotted arc represents the several cycles that can be performed while evaluating product increments and identifying new requirements.

The Sprint for identifying requirements (in terminology of SCRUM the identifying of the product backlog increments) follows the idea of the Human-Centred Design Process. Performing this process provides a good chance to develop software that has the necessary utilities, is usable and provides the necessary user experience.

Figure 8 provides more detailed information of activities that have to be performed according to the HCD process.

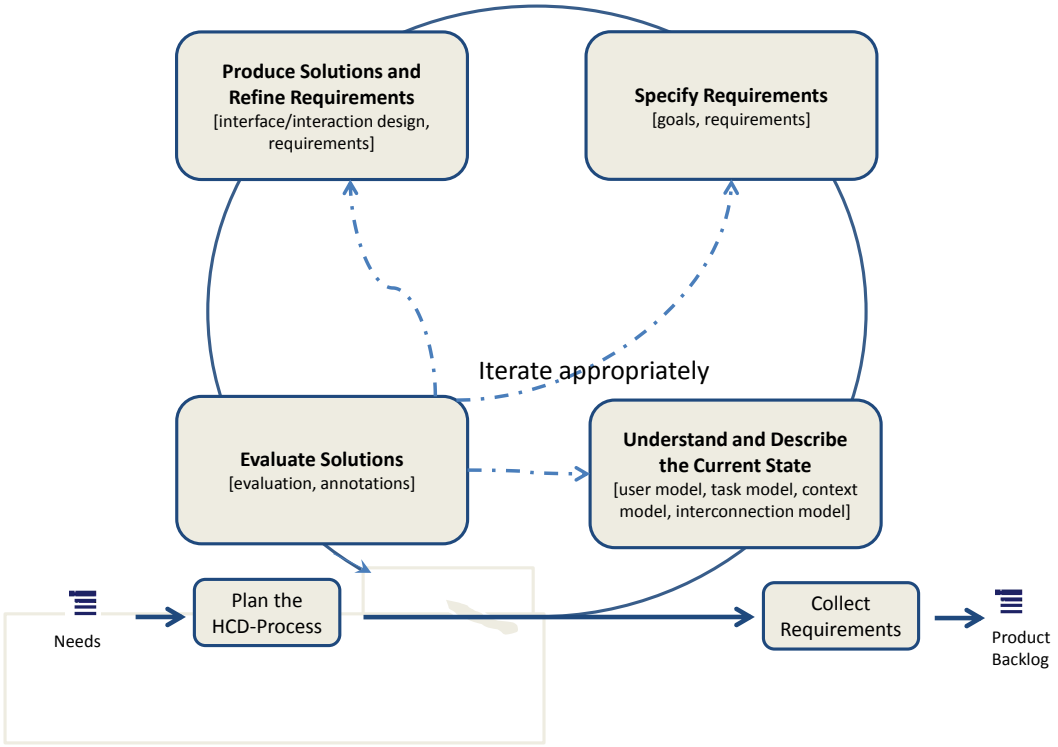


Figure 8. Detailed Human-Centred Design Process of Figure 7

Based on the identified needs, there has to be a plan for the activities in the forthcoming HCD-Process, that allows analysts identifying project

requirements under development. The current state and context of use is analysed and represented by different models. Requirements are specified afterwards before possible solutions are produced and later evaluated. The identified requirements are collected and specified as product backlog.

We do not focus on the SCRUM development cycle because our process model is independent from details of these models. More or less each agile development cycle can be supported.

4 Summary and Outlook

The paper discussed the two widely accepted process models for Human-Centred Design and Agile Development and shows how to combine them beneficially. Additionally, there was a focus on patterns that were published in different scientific papers. They encapsulate some knowledge about the integration of usability aspects in agile software development and provided arguments for presenting an own extended process model for the integrated approach of Human-Centred Design and SCRUM. Although the terminology of SCRUM was used the process model can be used in conjunction with any other agile development process. The provided solution is based on the *"Zero Sprint Pattern"*, *"One Sprint Ahead Pattern"* and the *"Parallel Track Pattern"*. It can be combined with the idea of having UX specialist as product owners and having UX specialists in the development team.

It might make sense to study the organisation of teams with the specific roles in more detail in the future. One consequence could be that all people acting in the development process should take part in the process of identifying needs on the project's vision.

References

[1] Agile Manifesto, http://agilemanifesto.org/, last visited June 4th 2015.
[2] Bertholdo, A.P.O., da Silva, T.S., de Melo, C.O., Kon, F. and Silveira, M.S. Agile Usability Patterns for UCD Early Stages, In Marcus A. (Ed.) DUXU 2014 Part I; LNCS 8517, pp. 33 – 44.
[3] Gamma, E., Helm, R., Johnson, R. and Vlissides, J. Design Patterns: Elements of Reusable Object-Oriented Software, Addison-Wesley Professional, 1994

[4] Herczeg, M., Kammler, M., Mentler, T. and Roenspieß, A. The Usability Engineering Repository UsER for the Development of Task- and Event-based Human-Machine-Interfaces, In: Narayanan, S. (Ed.). 12th IFAC, IFIP, IFORS, IEA Symposium on Analysis, Design and Evaluation of Human-Machine-Systems. Las Vegas : International Federation of Automatic Control. pp. 483-490

[5] Memmel, T., Gundelsweiler, F. and Reiterer, H. 2007. Agile human-centered software engineering. In Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI...but not as we know it - Volume 1 (BCS-HCI '07), Vol. 1. British Computer Society, Swinton, UK, UK, 167-175.

[6] Paelke, V. and Nebe, K. Integrating Agile Methods for Mixed Reality Design Space Exploration. In Proceedings of the 7th ACM conference on Designing interactive systems (DIS '08). ACM, New York, NY, USA, 240-249. http://doi.acm.org/10.1145/1394445.1394471

[7] Paul, M., Roenspieß A., Mentler T., Herczeg M. (2014). The Usability Engineering Repository (UsER). In Hasselbring, W & Ehmke, N C (Eds.) Software Engineering 2014 - Fachtagung des GI-Fachbereichs Softwaretechnik, 25.-28. Februar 2014, Kiel. Gesellschaft für Informatik e.V. (GI). 113-118

[8] Paul, M. Systemgestützte Integration des Usability-Engineerings in den Software-Entwicklungsprozess, PhD Thesis, Univerity of Lübeck, 2015.

[9] Rising, L. and Janoff, S. N. "The Scrum Software Development Process for Small Teams", IEEE Software, vol.17, no. 4, pp. 26-32, July/August 2000, doi:10.1109/52.854065

[10] Salah, D., Paige, R. and Cairns, P. A Practitioner Perspective on Integrating Agile and User Centred Design, Proceedings of the 28th International BCS Human Computer Interaction Conference (HCI 2014),

[11] Singh, M. U-SCRUM: An agile methodology for promoting usability, Integrating usability engineering and agile software development: A literature review. In Proc. AGILE 2009, IEEE Press, pp. 555-560.

[12] Sohaib, O. and Khan, K, Integrating usability engineering and agile software development: A literature review. In Proc. International Conference on Computer design and Applications (ICCDA), Volume 2, p. 32 – 38, 2010.

[13] Sutherland, J. Harrison, N. and Riddle, J. Teams That Finish Early Accelerate Faster: A Pattern Language for High Performing Scrum Teams. In Proc. HICSS 2014, pp. 4722-4728.

Continuous User Experience Development

Kati Kuusinen
Tampere University of Technology
Tampere, Finland
Korkeakoulunkatu 1, FI-33101 Tampere
kati.kuusinen@tut.fi

Abstract. Continuous approaches for software engineering such as continuous planning, development, and operations are becoming increasingly popular in agile software companies. It means that also user-centered design practitioners and practices need to adapt to both possibilities and challenges the increasingly rapid and more tightly integrated software engineering environment induces. Such issues include planning continuously throughout the life cycle instead of upfront planning, delivering user value whenever something is ready instead of delivering working software at the end of time-boxed iterations, and experimenting with real users instead of conducting traditional user studies and tests. In this position paper we discuss how user experience work can be organized with continuous software development.

1 Introduction

Continuous software engineering activities are means to enable rapid development and release cycles in companies. They seek to interweave activities that traditionally have been disconnected in the software lifecycle such as planning, design, and implementation [5]. Continuous software engineering activities often combine automation in software development and deployment processes in a way that minimizes the time required for such practices as integration, verification, deployment and delivery of software [5]. In addition to automated practices, continuous can also refer to practices conducted manually on a steady basis [5]. Fitzgerald et al. [5] define that continuous development consists of the following continuous activities: integration, deployment, delivery, verification and testing, security, and compliance. In addition to plain software development activities, the entire software life cycle can include continuous activities related to planning, operations, and improvement [5].

In relation to agile software development methodologies, the concept of *continuous* can be equated with the concept of *flow* used in Lean development [5], and continuous integration is a practice utilized in Extreme Programming [1]. Bosch [2] claims that continuous integration has no business value if the

organization does not follow agile working practices. Continuous development practices can be utilized with any agile development methodology. However, as there is a constant possibility to deliver whenever something is ready, continuous delivery can diminish the idea of time-boxed iterations utilized, for instance, in Scrum.

In the context of user-centered design (UCD) and user experience (UX) work, continuous software engineering offers benefits such as easier to arrange user experiments. As the development pipeline offers easier deployment of features whenever they are ready for user testing or production, features can be exposed to real usage within a target user group. In case of failure or, for example, after an A/B testing period, features can be withdrawn as easily from the system. Thus, feedback from real users' actual usage can be collected more rapidly and easily than with conventional methods. On the other hand, there is less time for keeping UX design activities ahead of development in continuous software development.

In this paper, we discuss ways to realize UX work in continuous software engineering. We call the approach continuous UX development (CUD). Our focus is in aligning the practices of UX work and software development in a way that forms a coherent interplay between the disciplines allowing UX specialists and software developers to work as a single cross-functional team.

2 Previous Research on Agile UCD

According to a recent systematic literature review [3], current approaches for agile UCD emphasize the role of keeping the UX design work ahead of development, often by iteration or more. In addition, most of the research on agile UCD integration considers it necessary to conduct some design upfront work before starting agile development iterations [3]. Sy [7] has introduced the most commonly referred framework that implements both of those principles. Benefits of such approach include that it offers UX designers time for user studies and design work with less time pressure from the developer side. However, the approach has also its limitations. Despite of its popularity, research conducted in companies utilizing the one iteration ahead approach report challenges in deciding on the scope of design upfront work, in chunking the UX design to suit agile iterations, and in synchronizing the UX design and software development tasks during the iterations [6]. In addition, the approach decreases agility of the project since the team needs to plan the work at least two iterations ahead. Also, since the UX design is created in advance, it might be more prone to changing conditions that can lead to design waste. Moreover, there is some evidence that an approach of UX designers and developers

working as a single cross-functional team might be more efficient than the one iteration ahead approach [4].

3 Continuous UX Development

Since continuous development practices aim at mitigating boundaries between different activities and phases of software life-cycle and at delivering whenever something is ready, we see it necessary to involve the UX specialist in the development team itself. In this setting, UX design work should be conducted in the cross-functional development team on as-needed basis; the one iteration ahead approach is not feasible with continuous development. Next, we describe how we see continuous UX development (CUD).

When a project is initiated, a mini-team starts to work towards a minimum viable working prototype of the software. The mini-team can consist of, for example, one UX specialist and one to two developers. The team together with users works towards understanding the project vision and most critical user paths. This can be done in, for instance, repeated workshops. Between the workshops the UX specialist designs the user flow and drafts user interface that are iterated together with the users in following workshops. Simultaneously, the developer(s) build up the technology stack and the pipeline for continuous development, and experiment with possible technologies they will utilize in the project. In addition, the team, together with the product owner, creates the initial backlog. A minimum viable prototype is an early version of minimum viable product that realizes the design idea. The minimum viable prototype can be, for example, a clickable software prototype with simulated backend that realizes user paths of most important use cases. We believe that a software prototype that allows actual user interaction is the best for communicating the design idea for the user.

After the minimum viable prototype is validated with users, the team size is readjusted to meet the capacity requirements of the project. The team starts to work towards the first production version of the system. The UX specialist either implements the user interaction or pairs up with a front-end developer. The user interaction is built based on continuous communication within the team and together with users when needed. When the team cannot solve a UX issue, developers start to build the next task on the priority list and the UX specialist investigates the problem until a solution is found. The team continues with similar setting during forthcoming iterations.

In addition to working software, we recommend allowing the delivery of working prototype and partially functioning features for gaining user feedback. That means the system can contain both fully working features and

forthcoming features that are delivered for some user groups in order to allow getting early user feedback before launching the feature. This approach is especially beneficial for situations where actual A/B testing is not applicable due to smaller number of users.

The CUD approach we suggest in this position paper offers an alternative to the commonly recommended one iteration ahead approach. Our aim is to offer a way to minimize the required amount of user studies and instead focus on building the user interaction modularly feature by feature. The approach necessitates that the cooperative development team includes a UX specialist. Moreover, it requires a mindset that allows trial and error: when the UX design requires alterations, the user interface will be iterated and refactored. Compared to the one iteration ahead approach, we expect CUD to offer better visibility to the common vision, or the big picture, of the project. Moreover, we expect CUD to improve the team communication and increase developers' commitment towards UX tasks. In addition, we expect to get feedback from actual usage faster than in the one iteration ahead approach. While user studies are valuable, it is the actual use that really validates the viability of the system.

4 Background and Framing of the Position Paper

We base this position paper on our studies of UX development in various software companies. However, we have seen the approach in use as such in none of those companies. Our background is in studying the development of enterprise software and work-related tools. Thus, the approach we present here has been influenced by development of practical-oriented software tools. Therefore, to be transferable to leisure system development, such as game development, the approach may require some changes. First, the focus on tasks directs towards defined user paths. Second, it is usually known in work-related software development who the users will be. Moreover, enterprise software is often built to solve a current business problem of the user organization. Finally, the UX work in enterprise software often focuses on fluent user interaction and practical quality of use whereas focus of leisure systems development can be more in the hedonic. Thus, we believe that UX work of leisure systems should involve more robust methods for evoking users' emotions and hedonic experiences. In addition to leisure systems, considerations related to scalability and special application areas such as safety-critical remain future work.

5 Conclusions

This position paper introduces an alternative to the popular one iteration ahead approach for integrating user-centered design with agile software development methodologies. Moreover, the approach is compatible with continuous software engineering practices. We suggest to involve a UX specialist as a member of the development team, and to conduct UX work in the same iteration together with development practices. To succeed with such approach, the whole team, together with the user, needs to cooperate on daily basis. Furthermore, upfront studying and planning should be minimized and instead trial and error in design and development activities should be allowed.

References

[1] Beck, K. Extreme Programming Explained: Embrace Change. Addison-Wesley, 1999. 224p.

[2] Bosch, J. Continuous Software Engineering: An Introduction. In J. Bosch (ed.), Continuous Software Engineering, Springer International Publishing Switzerland, 2014. pp. 3-13.

[3] Brhel, M., Meth, H., Maedche, A., and Werder, C. Exploring principles of user-centered agile software development: A literature review. Information and Software Technology, 61, 2015. pp. 163-181.

[4] Ferreira, J., Sharp, H., Robinson, H. User experience design and agile development: managing cooperation through articulation work. Software: Practice and Experience 41, 9, 2011, John Wiley & Sons. pp. 963-974.

[5] Fitzgerald, B., and Stol, K.-J. Continuous software engineering and beyond: trends and challenges. In Proc. 1st International Workshop on Rapid Continuous Software Engineering – RCoSE 2014. ACM, New York, NY, USA, 2014. pp. 1-9.

[6] Salah, D., Paige, R. and Cairns, P. A systematic literature review on agile development processes and user centred design integration. Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering – EASE'14. ACM, 2014. Article 5, 10 p.

[7] Sy, D. Adapting usability investigations for Agile user-centered design. Journal of Usability Studies 2, 3, 2007. pp. 112–132.

Addressing Usability and UX in Call for Tender for IT Products

Rosa Lanzilotti, Maria Francesca Costabile, Carmelo Ardito
Dipartimento di Informatica, Università degli Studi di Bari Aldo Moro
via Orabona, 4 – 70125 Bari (Italy)
name.surname@uniba.it

Abstract. A body of research in literature has demonstrated the benefits for software companies of adopting methods to address usability and user experience. However, as revealed by the very poor usability of many software systems, companies very often do not consider such methods in their development practices. From our previous studies it emerged that one of the reason is that usability and UX are quality requirements not formally established in the Call for Tender (CfT) for developing software products to which they respond. Therefore, they do not commit resources to satisfy usability and UX requirements. In this paper, we report a systematic review of 44 national and international CfTs for IT products, which aimed at investigating whether and how usability requirements are mentioned.

1 Introduction

Since the 80s, when Human-Centred Design (HCD) was proposed, a significant amount of research has been done with the aim to define methods that support professionals, who work in companies that produce software, in designing and evaluating interactive systems. However, our experience confirms that HCD and other methods to ensure usability and good user experience (UX), even if always mentioned in the literature, are very rarely applied in the actual practice of software development [3, 5, 12]. Today many companies invest most of their resources to produce software that can be competitive on the market. However, the daily experience with a number of software products show that, despite the powerful features offered, users experience many difficulties in using them. A user interface that is difficult to understand and use causes many problems. Usability is a measure of the degree to which users are able to conduct their activities in their specific context of use with effectiveness, efficiency and satisfaction [11]. A low level of usability means that users cannot understand how to use the interactive system, regardless of its functional complexity [13]. Therefore, usability becomes the main goal and usability evaluation represents a fundamental activity of the entire software development.

We increasingly feel the need that this situation must be changed. Several researchers have suggested methods and techniques to maximize the impact of

usability and UX of software products. Our experience with companies has shown that presenting to them methodologies and techniques to be used in their practice of software development it is not enough. We realized that the real integration of activities aimed at improving the usability and UX products requires a thorough analysis of the software development practices, performed together with the professionals while working with them within the company. Through this collaboration, methodologies and techniques for the assessment of usability and UX can be identified and adapted to the specific company.

In recent years, researchers from the IVU (Interaction, Visualization, Usability and user experience) laboratory of the University of Bari Aldo Moro have spent a considerable effort to transfer in companies and organizations methods and tools for dealing with UX along their software development practices [2-4]. More recently, an experimental study has been conducted with the aim to investigate to what extent companies take into account usability and UX, and to analyse how they can modify their process of development, in order to create better products. It represents one of the few studies conducted with the involvement of software companies, in order to reduce the gap between what is proposed by the academic world - and widely published in the literature - and the actual practices of software development. The results show that still too many companies overlook important quality factors, such as usability and UX. The study also revealed an important reason why companies do not take into account usability engineering methods. In their development processes companies focus primarily on the requirements formally established in the Call for Tender (CfT) to which they respond. In general, these requirements do not include usability and, therefore, they do not commit resources to satisfy usability and UX requirements. For example, one respondent said: "*I do not burden my company with what it is not explicitly required in the call for tender*."

As described in Section 2, in order to investigate this result, we performed a systematic review of 44 national and international CfTs for IT products, with the aim to investigate whether and how usability requirements are mentioned. In Section 3, we report the results of an online questionnaire that we also administered to experts in the field of ICT, in order to collect their opinions about the presence of usability requirements in call CfTs for IT products. Section 4 concludes the paper and outlines future work.

2 Usability requirements in CfTs for IT products

In the last few months, we have performed an analysis of 44 CfTs for IT products issued by public and private organizations (26 in Italy and 18 in

European countries) in order to investigate to which extent the CfTs explicitly indicate usability and UX requirements. The public organizations that issued the analysed CfTs ranged from National Government institutions (e.g., Ministry of University and Research, Ministry of the Interior) to small municipalities. The private organizations were Small-Medium Enterprises. Preliminary results revealed five different categories (see Figure 1).

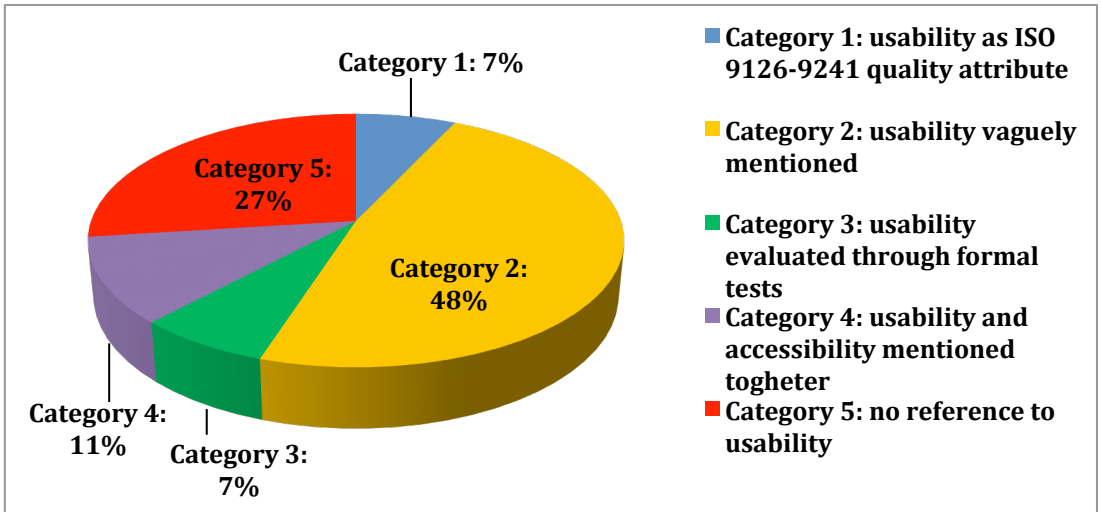


Figure 1. Results of the analysis to investigate to which extent 44 CfTs for ICT systems issued by public organization explicitly indicate usability and UX requirements.

Category 1) The 7% (N = 3) of the analysed CfTs only mention usability by referring to the ISO 9241 and 9126 standards [10, 11]. For example, in the CfT of the Italian Apulia Region government for the development of a system for registering employees presence there is only a sentence that refers to usability [6]: "*Application programs should preferably meet the quality requirements, as those reported in the ISO 9126 (i.e, functionality, reliability, usability, efficiency, maintainability and portability)*". The word "preferably" clearly indicates that software quality is not mandatory, in particular no specific requirements about usability are provided, so it is not clear if and how companies would consider it.

Category 2) In the 48% (N = 21) of the analysed CfTs, more attention is devoted to usability aspects, but they are still expressed in very general terms. For example in an Italian CfT [7], it is written "*The website must be characterized by its immediacy [i.e. learnability] and ease of use*", "*It must be organized clearly [...]*", "*Great care must be devoted to creativity, ..., usability and accessibility*". Again, it is not clearly specified how the usability level of the final products has to be measured.

Category 3) The 7% (N = 3) of the calls explicitly require to perform formal usability tests. For example, in the CfTs issued by the International

Centre for Trade and Sustainable Development (ICTSD) for the redesign of their website [8] it is written: "*Usability: Clarity of interaction between the users and the site is crucial, and must be paramount in the redesign. The web team will run at least two usability tests throughout the design and development process of the new site (to create a short list of the most serious problems and a commitment to fixing them before the next round of testing). A specific usability testing plan (with details on number of participants in each round, who we test with, where/when we test, who watches, reporting, etc.) will be released before the beginning of the design process*." Thus, a rigorous process to verify the usability of the software system is required. Two usability tests should be conducted, with particular attention to subjects sampling, tasks execution and results reporting. The final report of the first test will identify a list of the most serious problems, which have to be fixed before running the second usability test.

Category 4) In the 11% (N = 5) of the analysed CfTs, the concept of usability is mixed with that of accessibility. For example, in [9] it is reported: "*Most importantly, they should demonstrate their understanding of and ability to implement accessibility and usability standards and maintain a high level of practical accessibility*". In some countries, like in Italy, there are laws that clearly regulate the levels of accessibility of software products. For example, in Italy it is mandatory to comply with the Stanca law about accessibility, which was approved by the Italian Parliament in 2003. Thanks to this law, it is easy to specify accessibility requirements in the CfTs. As a consequence, attention is devoted to product accessibility but not to usability.

Category 5) The 27% (N = 12) of the analysed CfTs did not mention usability at all.

3 Online questionnaire on usability in CfTs

An online questionnaire was administered to experts in the field of IT, in order to collect their opinions about the presence of usability requirements in CfTs for IT products. The questionnaire was created and made available online using Google Forms. Various social media, such as Facebook, were used to promote and share the questionnaire to different groups, including Italian Startup Scene Italian Startup Scene - Puglia, and Digital Natives.

The questionnaire was structured to address:

- • *Respondents' profile*: to identify the basic characteristics of surveyed users, i.e. socio-demographic, academic qualifications, work experience.
- • *Experience with call for tender*: to check the presence of usability requirements in CfTs, to collect information on metrics and guidelines

possibly specified in the calls, and to know the respondents' opinion about advantages/disadvantages of including usability requirements in CfTs.

In the following we summarize the collected data.

3.1 Respondents' profile

The questionnaire was completed by 54 professionals and / or experts in the ICT sector. 44% of the respondents (hereinafter also referred to as participants) were developers, 15% designers, 11% managers, 7% analysts, 22% other people (data analysts, computer experts, students, consultants, systems analysts, and so on). 46% is made up of first-level graduates, 28% of second-level graduates, 22% of high-school graduates, 2% of PhDs, 2% do not indicate any qualification.

Regarding work experience, 46% of participants have between 1 and 5 year experience, 32% have been working for more than 10 years, 11% have between 5 and 10 years of experience and another 11% have less than 1 year experience.

91% of participants claimed to be familiar with usability. However, when asked to provide a definition of usability, only 15% of them gave a correct definition, 33% a definition partially correct; 6% indicated references where to get a usability definition. The answers of 28% participants revealed that they confuse usability and accessibility; 7% gave an incorrect definition, 11% clearly said that they are completely unaware of usability.

Another important result concerned the deployment of usability evaluations in the participants' companies. 35% of the participants confirmed the presence of a usability expert in their company and indicated interviews, questionnaires, focus groups, and direct observation, as the methods mostly used. The remaining 65% of the participants said that their company has no usability experts.

3.2 Experience with call for tender

The first question asked to indicate whether the participant had read at least once a CfT. 70% responded positively. We first asked them how frequently they found usability requirements clearly expressed in the CfTs. 60% said rarely, 24% never, 13% often, and only 3% said they had found usability requirements in each call they have read (see Figure 2). 90% of those who have found usability requirements in CfTs have never found explicit indications about methods or guidelines, but only references to documents issued by CNIPA (Centro Nazionale per l'Informatica nella Pubblica

Amministrazione, or, in English, National Centre for IT in Public Administration), DigitPA (Digital Public Administration) and AgID (Agenzia per l'Italia Digitale, or, in English, Agency for Digital Italy).

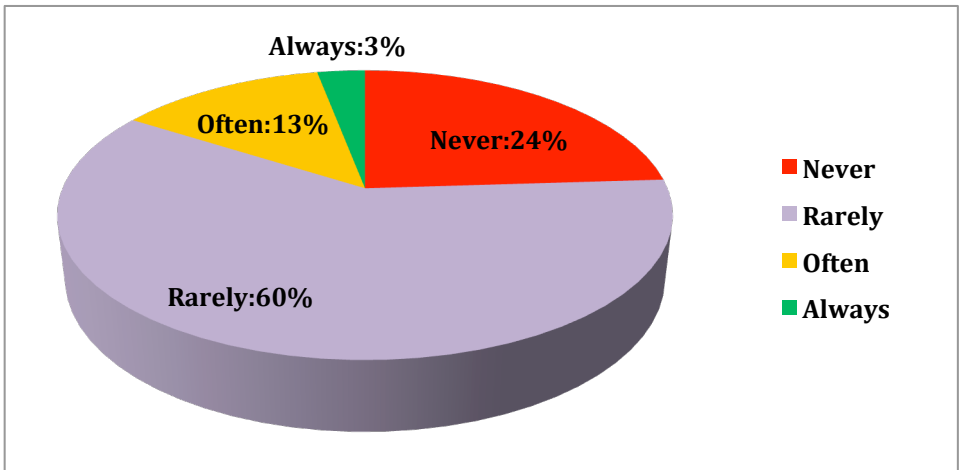


Figure 2. Presence in the CfTs of usability requirements clearly expressed, according to interviewees' experience.

Then respondents were asked to give their opinion on the convenience of including usability requirements in CfTs, also indicating advantages and disadvantages. 89% of respondents agreed on the need to include usability requirements in CfTs. In fact, they believe that this would bring benefits to end users in terms of ease of use of the final product, reducing the time spent to request support and assistance, decreasing the number of mistakes, improving productivity, etc. Also software companies can beneficiate, because of the lesser costs for system maintenance, increased productivity, improvement in the ROI (Return On Investments). The remaining 11% partially or totally disagreed with the inclusion of usability requirements in CfTs. 50% of them believe that the inclusion of usability requirements would cause an increase in production costs; 33% thinks that this would increase the resources for carrying out usability evaluations; 17% believes that there are other reasons why usability requirements should not be in CfTs. For example, "*There is no an objective way to evaluate usability, so it must stay out of CfTs. The only way to evaluate it is through controlled experiments on a sample, and it is extremely expensive if you want that the sample is even vaguely representative of the final users. In any case it is difficult to generalize the results*." Such a response highlights the huge confusion of this participant about usability evaluations.

4 Conclusion and future work

The analysis of 44 national and international CfTs showed that many of them do not mention usability requirements or refer to such requirements very vaguely. As a consequence, companies do not bother to develop their products following a Human-Centred Design approach, because their goal is to satisfy the requirements clearly expressed in the CfTs. To change the current situation, it is necessary to convince the public and/or private organizations to mention explicitly the usability requirements in their calls, so companies will be obliged to consider them. On the other hand, the research community must work to define usability requirements which are objectively measurable.

We are currently collaborating with members of the GLU (Gruppo di Lavoro per l'Usabilità), an Italian working group on usability related to the Italian Ministry of Public Administration. The GLU aims at improving the usability of Public Administration websites and other e-government systems. Indeed, in May 2015, the GLU published a new version of a document that provides detailed guidelines for the design of the websites of the Public Administration, called eGLU Protocol 2.1 [1], whose aim is to guide web masters and web editors, who do not have experience on usability and UX evaluation, in the identification of usability problems of the websites they work on, by committing very few resources in terms of time and people. This document describes all the needed tools (e.g. modules to report the usability problems, spreadsheets to analyse the acquired data according to the identified measurement criteria) and the specific steps to organize and perform an effective usability test by using the thinking aloud technique, which is well known for allowing to perform an accurate evaluation at low cost [13]. Together with the GLU members we are also working at the definition of a framework to facilitate the inclusion and specification in CfT of requirements related to the adoption of HCD techniques by software companies. Depending on the complexity of the system to be developed, the framework prescribes to adopt different HCD techniques (e.g., personas, scenarios, interviews, questionnaires, user tests).

Acknowledgements

This work is partially supported by the Italian Ministry of University and Research (MIUR) under grants PON 02_00563_3470993 "VINCENTE" and PON04a2_B "EDOC@WORK3.0", and by the Italian Ministry of Economic Development (MISE) under grant PON Industria 2015 MI01_00294 "LOGIN".

References

[1] eGLU. http://www.funzionepubblica.gov.it/glu.aspx. Last access: June 4th, 2015

[2] Ardito, C., Buono, P., Caivano, D., Costabile, M. F., Lanzilotti, R., Bruun, A. and Stage, J. 2011. Usability evaluation: a survey of software development organizations. In Proceedings of the International Conference on Software Engineering and Knowledge Engineering. SEKE '11. Knowledge Systems Institute, Skokie, 282-287.

[3] Ardito, C., Buono, P., Caivano, D., Costabile, M. F. and Lanzilotti, R. 2014. Investigating and promoting UX practice in industry: An experimental study. International Journal of Human-Computer Studies, 72, 6 (2014), 542-551.

[4] Ardito, C., Buono, P., Caivano, D., Costabile, M. F., Lanzilotti, R. and Dittrich, Y. 2014. HCI and SE Practices in Industry: Lessons from the Trench. IEEE Computer, 47, 12 (Dec. 2014), 86-89.

[5] Cajander, A., Gulliksen, J. and Boivie, I. Management perspectives on usability in a public authority: a case study. Oslo, Norway, 2006. ACM, New York, NY, USA.

[6] Call for tender - Sistema di rilevazione presenze del personale della Regione Puglia. http://www.regione.puglia.it/index.php?page=burp&opz=getfile&file=17.htm&anno=xl&num=166. Last access: June 4th, 2015

[7] Call for tender - Procedura aperta per la progettazione, realizzazione, manutenzione, la redazione dei contenuti e le azioni di comunicazione social del sito web istituzionale dell'Autorità garante per l'infanzia e l'adolescenza. http://www.garanteinfanzia.org/sites/default/files/documenti/capitolato%20di%20gara%20sito%20web.pdf. Last access: June 4th, 2015

[8] Call for tender - Redesign the Corporate Website. http://ictsd.org/redesign/wp-content/uploads/2013/09/i-Call-for-Tenders_Design.pdf. Last access: June 4th, 2015

[9] Call for tender - Development and provision on the web of a digital product/service in favour of citizens, operators and entities participating in the trial test in Rome area. http://www.crossproject.eu/wp-content/uploads/2013/11/Call-for-tender-APP4CROSS.pdf. Last access: June 4th, 2015

[10] ISO/IEC 9126 1991. Information technology - Software Product Evaluation - Quality characteristics and guidelines for their use. Available at, Last access.

[11] ISO/IEC 9241-11 1998. Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability. Available at, Last access.

[12] Larusdottir, M. K. 2012. User Centred Evaluation in Experimental and Practical Settings. PhD thesys, KTH Royal Institute of Technology, Stockholm, Sweden.
[13] Nielsen, J. 1993. Usability Engineering. Morgan Kaufmann.

User and Client Satisfaction in Agile Development

Marta Larusdottir[1], Effie Law[2], Åsa Cajander[3]
[1]School of Computer Science, Reykjavik University, Iceland,
Menntavegur 1, 101 Reykjavik
[2] Department of Computer Science, University of Leicester, UK,
LE1 7RH, University Road, Leicester
[3]Department of Information Technology, Uppsala University, Sweden,
Box 337, SE-751 Uppsala
[1]marta@ru.is; [2]lcl9@leicester.ac.uk;
[3]asa.cajander@it.uu.se

Abstract: The emphasis on User Centred Design (UCD) in agile systems development processes (Agile) has been studied from various perspectives. The context of Agile strongly affects the possibilities for IT professionals to conduct user-centred activities in their work. In this paper, we describe a survey study comparing the responses from professionals using Agile processes and other software development processes. Specifically, we explore the values and perspectives that professionals emphasise in their work, whether feedback is gathered from stakeholders, and how frequently feedback is gathered. The main results show that both user and client satisfaction is emphasised by professionals using Agile, but for professionals using other processes such as their home-grown process or the traditional waterfall approach, the focus is user satisfaction. The survey involved a relatively small number of participants (N=42), and thus can be seen as an exploratory work that can inform our future work.

1 Introduction

According to ISO/IEC 12207: 2008 [10], the goal of software development is to develop a set of computer programs (software), procedures and associated documentation and data. During software development, there are various phases, starting from the acquisition of the software, to the supply, development, operation, maintenance, and disposal of the software. Depending on the type and scale of a project (e.g., in-house vs. external organization order; short-term vs. long-term; local vs. global market), which in turn influences the choice of software development process and methods, different stakeholders can be involved. They include project manager, client, customer, software analyst, software developer, interface designer, usability and user experience specialist, and, presumably the most important among all, **end-user.** This list as well as the characteristics (e.g., role, expertise) of individual stakeholders can further be refined contingent on a project's profile. The

emphasis on fulfilling the needs of the different stakeholders varies with the software development process used.

Agile software development processes (Agile), such as Scrum, Extreme Programming (XP) and Dynamic Systems Development Method (DSDM), have become a *de facto* standard for software development practice. Lean software development, which is part of the Agile category, has recently gained popularity, with the Kanban processes being the most popular one [11]. Agile focuses on delivering functioning software early and continuously. Agile values speed, communication and collaboration in software development. In the Agile Manifesto, customer collaboration is one of the four core values and the first principle is described as "*our highest priority is to satisfy the customer through early and continuous delivery of valuable software"*, so customer satisfaction is strongly recommended [2]. In ISO/IEC 12207: 2008 [10], definitions of different stakeholders involved in software development, including customer, are given. Accordingly, customer is defined as: "*organization or person that receives a product or service* and a user is defined as: *individual or group that benefits from a system during its utilization"*.

In accordance with ISO 9241-210: 2010 [9], the User Centred Design (UCD) framework involves several key approaches: user-centred systems design methods, rapid contextual design, and user involvement. In UCD, we address different aspects of user experience (UX) and usability of software with user satisfaction being a major focus. In ISO 9241-210: 2010 [9], a user is defined as "*a person who interacts with the product"*. Examples of UCD activities include creating personas to communicate user research, doing field studies that observe users, and usability evaluations for gathering user feedback.

The recent wide adoption of Agile processes can be explained by the implicit assumptions that they address user perspectives better than traditional software processes [1] and that by simply applying an Agile development process the software system can become more usable for end-users than otherwise. However, previous research has shown that this is not always the case, and that the context of Scrum impacts user involvement as described in, for example [4] and [12]. Many research projects aim to analyse and understand the conditions under which Agile and UCD may work together. Some studies are based on empirical data (e.g., [8], [14]), but some are rather analytical, expressing opinions based on certain ideas and assumptions, which are thus more open to debate.

Another intriguing topic is the integration of UCD within Agile. Chamberlain and colleagues [5] conducted a field study to investigate the

integration of UCD into Agile. They concluded that successful integration requires *balancing* between each of the disciplines in the team and that sufficient resources for the work need to be provided. Additionally, it requires that all key members must be involved in key decision points in the project and that users play an important part in the project. Blomkvist [3] claimed that Agile processes do not inherently provide the required support for including user perspectives in development. As an example, iterative development is fundamental to both UCD and Agile, but their views on as well as definition of the term "*iterative"* are substantially different [3].

Nevertheless, the basic values and specific methods of Agile may have the potential to work well together with UCD. Ferreira and colleagues [8] conducted an observational study of a mature Scrum team in a large organization, and their interactions with the UX designers working on the same project. They concluded that the cooperation between Agile developers and UX designers was achieved through on-going *articulation work* by the developers, who were compelled to engage a culturally distinct UX design division. Constantine ([6], [7]) reworked his Usage-centered design methodology to become more lightweight. He claimed that his model-based approach focused on usability and user interface design, and even more so when it was turned into an Agile version of the same overall methodology. McInerney and Maurer [13] interviewed three usability specialists in Agile projects. They were all very positive over their ability to manage usability and UX activities in the Agile projects, and although they could not prove any positive effects in the resulting projects, they were positive about their ability to contribute, and did not identify any negative effects resulting from the adoption of the Agile approach in the projects [13].

In this paper, we address three particular questions to explore if there are any differences in attitude and practice between software development professionals who use agile processes and those who use other processes:

a) What do professionals state as the main emphasis or values when developing software?
b) Do they gather feedback from various stakeholders?
c) If they gather feedback, how frequently is that done?

2 Method

In this section we describe the method and procedure of the study, the participants in the survey, the processes they used for software development and their main job role.

2.1 Method and Procedure

The survey consisted of 48 questions, 40 close-ended and 8 open-ended. The first part of the survey collected information regarding the participants' background, work environment and experience. The second part contained 4 questions regarding the participants' use of software development processes and their preferences. The third part contained a combination of 14 open-ended and close-ended questions to gather their understanding of the terms 'user', 'customer' and 'client' and to describe the people belonging to these three groups in their development project. Additionally, the participants were asked to describe their main emphasis when developing software. Finally, they were asked to respond to the questions about gathering feedback on design artefacts from stakeholders, including users, customers, clients, and colleagues. In this paper, we focus on the analysis of the participants' responses to the questions on the main emphasis and gathering feedback from stakeholders.

2.2 Participants

The survey was web-based and distributed via email to 393 graduates from Computer Science of Reykjavik University in Iceland that all had successfully completed at least a B.Sc. program there and their graduation year was between 2009 and 2014. Out of these graduates, 73 responded to some questions in the survey (i.e. the response rate of 18.6%). To the questions analysed in this paper, 42 responded (i.e. the response rate of 10.7%), the gender distribution was 74% male, 12% female and 14% did not respond to that question. This response rate was rather low, but the study could be seen as a pilot study for future work in the area.

All of the 42 participants received the more or less the same training in computer science and all of them completed a course in Human-Computer Interaction (HCI) as a compulsory part of their education. Of the respondents, 86% completed a B.Sc. degree, 14% completed a M.Sc., or a Ph.D. degree. Their industrial experience varied, 36% had been working for less than a year, 40% for 1 – 3 years and 24% for more than a year. The largest group of the participants (57%), at the time the survey was administered, worked in a company that had less than 50 employees, 17% at a company with 51 to 200 employees and 26% with over 200 employees.

Concerning the types of software developed in the participants' companies, the sector "Business/Finance" and "Data Management" were the most common (26% each), followed by games (14%), communication software (7%) and other categories (27%), including management and monitoring

software, web development software, specialised software and research software.

2.3 Development process

The most frequently used software development process was Scrum with 40% using that process only and 14% using both Scrum and Kanban. Another 12% were using only Kanban and 10% were using Agile processes other than Kanban or Scrum. In total 76% of the participants were using Agile processes (N=32) and 24% using processes other than Agile (N=10), including 19% using their own process and the others using the waterfall process or no process. The results in the paper have been derived from the analysis of two groups of participants: those using Agile (N=32) and those using other processes (N=10).

2.4 Main Job Role

The participants were asked what their main job role in the last three months was; the following options were given and more than one could be chosen: a) requirement analysis, b) design, c) programming, d) testing, and e) other activities (such as project management, teaching and researching, consulting, etc.). About 83% of the participants selected programming as their main task in their workplace. Other options included design (53%, the second highest), requirement analysis (~31%) and software testing (~31%).

Table 1. Number of main job roles of the participants.

Main job role	Agile (N=32)		Other (N=10)	
One	6	19%	1	10%
Two	8	25%	3	30%
Three	10	31%	3	30%
Four	5	16%	2	20%
Five	3	9%	1	10%

As shown in Table 1, around half of the participants using Agile development had 3 to 5 main job roles and about 60% of the participant using other processes. This suggests that each software developer needs to have various competences for developing software, at least in the software industry in Iceland.

3 Results

In this section, we structure the results according to the three research questions:

a) What do professionals state as the main emphasis or values when developing software?
b) Do they gather feedback from various stakeholders?
c) If they gather feedback, how frequently is that done?

3.1 Placing Emphasis

The participants were asked: "*Where do you place the most emphasis in the projects you are working on now*?"

Table 2**.** The main emphasis in the software projects

Emphasis	Agile (N=32)		Other (N=10)	
Client satisfaction	11	34%	0	0%
User satisfaction	10	31%	5	50%
Customer satisfaction	2	6%	0	0%
User and client satisfaction	1	3%	0	0%
Saving development time and/or cost	1	3%	2	20%
Safety	0	0%	2	20%
Hard to specify	2	6%	0	0%
Other	5	16%	1	10%

As shown by the results presented in Table 2, those participants using Agile most frequently put emphasis on client satisfaction but user satisfaction was similarly emphasised, and one participant stated that he emphasised both user and client satisfaction. It is interesting to note that none of the participants using other processes mentioned client satisfaction, but half of them mentioned user satisfaction as their main emphasis.

3.2 Gathering Feedback

The participants were asked if they gathered feedback from users, clients, customers and colleagues or friends. The motivation for asking was to check if they were using a user-centred approach to software development, and if they also consulted other stakeholders for getting feedback on their software. The results are shown in Table 3. In the survey, there was an explanation of what gathering feedback means.

Table 3. Gathering feedback from stakeholders

Gathering feedback from:	Agile (N=32)		Other (N=10)	
Users	23	72%	6	60%
Clients	14	44%	3	30%
Customers	9	28%	4	40%
Colleagues and friends	22	69%	6	60%

As shown in Table 3, 72% of the participants using Agile gathered feedback from users and 69% from colleagues and friends. Sixty percent of participants using other processes than Agile gathered feedback from users and there were also 60% that gathered feedback from colleagues and friends. Not as many gathered feedback from clients and customers.

The participants were also asked about the definition of the terms. Around 40% of the participants using Agile processes thought that the term client was the same as the term customer, about 20% thought the term user was the same as the term customer, and similarly about 20% thought that the term user was the same as the term client. For the participants using processes other than Agile, 80% thought that the term client was the same as the term customer, 30% thought that the term user was the same as the term customer and similarly 30% thought that the term user was the same as the term client. It is interesting to note that the term customer clearly has different meanings in these two groups of participants.

3.3 Frequency of Feedback

The third issue we analyse in the paper is the question about how often the participants gather feedback from various stakeholders, including users, customers, clients, colleagues and friends.

Table 4. Frequency of gathering feedback from stakeholders

	Agile (N=32)				Other (N=10)			
Gather feedback from:	Once a week	1 - 3 times a month	2 -4 times a year	Never or missing	Once a week	1 - 3 times a month	2 -4 times a year	Never or missing
Users	7	9	5	11	3	1	1	5
	22%	28%	16%	34%	30%	10%	10%	50%

Clients	4	6	3	19	1	1	1	7
	13%	19%	9%	59%	10%	10%	10%	70%
Customers	4	2	3	23	3	0	0	7
	13%	6%	9%	72%	30%	0%	0%	70%
Colleagues and friends	16	5	1	10	3	1	2	4
	50%	16%	3%	31%	30%	10%	20%	40%

As shown in Table 4, the participants using Agile most frequently consulted colleagues or friends; 50% of them consulted users at least once a month, but only 20% of them consulted customers, despite the emphasis in Agile on doing so. The participants using other processes, the difference was not so clear, but there were only ten of them. In responding to the question what method was most useful for gathering feedback, the most frequently method used for gathering feedback from all stakeholders was meeting.

4 Implications

Based on the results above, we pose these questions that could be further discussed at the workshop:

- In Agile processes the focus seems to be both on the client and on the user, whereas in other software development processes the focus seems to be more on the user. Does this difference depend on the way the question has been asked?
- It seems clear from this survey that most systems developers have many different roles in systems development. This has implications for our computer science education and we need to prepare students for this kind of work. Do we prepare the students for this in an adequate way?
- Even though client satisfaction is of the top priority, feedback is not often gathered from clients. How can we address this problem?
- Half of the Agile software developers gathered feedback from users at least once a month, and the most useful method mentioned was meeting. Since meeting is the most frequently mentioned method and the most useful one: Should we place more emphasis on teaching how to conduct productive meetings?

Acknowledgements

We would like to thank Daniel Multykh, a BSc. student in Computer Science at Reykjavik University, who conducted the survey and gathered the data on which this paper is based.

References

[1] Baxter, G., & Sommerville, I. (2011). Socio-technical systems: From design methods to systems engineering. Interacting with Computers, 23(1), 4-17.
[2] Beck, K., Beedle, M., Van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., ... & Thomas, D. (2001). The Agile Manifesto.
[3] Blomkvist, S. (2006). User-centred design and agile development of IT systems (Doctoral dissertation, Department of Information Technology, Uppsala University).
[4] Cajander, A, Larusdottir, M.K., Gulliksen, J (2013) Existing but not Explicit - The User Perspective in Scrum Projects in Practice, INTERACT 2013.
[5] Chamberlain, S., Sharp, H., & Maiden, N. (2006). Towards a framework for integrating agile development and user-centred design. Extreme Programming and Agile Processes in Software Engineering, 143-153.
[6] Constantine, L. L. (2002). Process agility and software usability: Toward lightweight usage-centered design. Information Age, 8(8), 1-10.
[7] Constantine, L. L., & Lockwood, L. A. (2003, May). Usage-centered software engineering: an agile approach to integrating users, user interfaces, and usability into software engineering practice. Proc 25th Int Conf Software Engineering 746-747. IEEE Computer Society.
[8] Ferreira, J., Sharp, H., & Robinson, H. (2011). User experience design and agile development: managing cooperation through articulation work. Software: Practice and Experience, 41(9), 963-974.
[9] ISO 9241-210:2010. Ergonomics of human-system interaction -- Part 210: Human-centred design for interactive systems.
[10] ISO/IEC 12207:2008. *Systems and software engineering - Software life cycle processes*.
[11] Kniberg, H., & Skarin, M. (2010). Kanban and Scrum: Making the Most of Both. C4media.
[12] Lárusdottir, M., Cajander, Å., & Gulliksen, J. (2014). Informal feedback rather than performance measurements–user-centred evaluation in Scrum projects. Behaviour & Information Technology, Vol. 33(11), Vol 33 - issue 11, pg. 1118 – 1135.
[13] McInerney, P., & Maurer, F. (2005). UCD in agile projects: dream team or odd couple?. interactions, 12(6), 19-23.
[14] Plonka, L., Sharp, H., Gregory, P., Taylor, K. (2014) UX design in agile: a DSDM case study. XP conference 2014.

Engineering for User Experience: An Interactive TV Case Study

Michael M. Pirker [1], Regina Bernhaupt [1,2], François Manciet [2]

[1] Koestendorfer Str. 8, 5202 Neumarkt a.W., Austria
[2] IRIT, Université Paul Sabatier, 118 Route de Narbonne, 31062 Toulouse CEDEX 9, France
michael.pirker@ruwido.com,
regina.bernhaupt@ruwido.com,
francois.manciet@irit.fr

Abstract. This paper provides an experience report and position paper describing some subjective and personal insights we gained during the design and development of an interactive television application that had the enhancement of the overall user experience as the main target. Goal of this position paper is to point out practical experience that - as simple lessons learned - can help avoid that a project fails due to simple communication errors or role-specific different expectations.

1 Introduction

The claim to focus on the User Experience (UX) is widespread in the field of interactive television (iTV) [1]. Vendors in almost all European countries use to claim that their next generation of interactive TV or Internet Protocol TV will provide more content and a better experience when interacting with the TV. In the majority of cases, the claim for a better experience is closely linked to the introduction of a new interaction technique. For example, France, one of the most vivid IPTV markets, saw the introduction of (1) the usage of pointing (point-and-click) to support Web browser navigation on the TV [3], (2) the usage of gesture to perform basic control tasks like volume or channel change [9] and announcements for touch to be included in the next generation [10].

Applying a user and UX centered design and engineering process is one of the most important pillars that modern product development is based on. The development process for interactive TV products and applications should be based on a thorough analysis of the targeted users, their tasks, as well as the environment and context that the iTV or IPTV product or application will be used in [4]. Goal is to gather and analyse requirements, to develop alternative designs based on the gathered data, and to support iterative development and the evaluation of various prototypes with different qualities [7], [13]. Prototypes can range from paper prototypes, sketching the basic idea of the

application, to final or nearly final products, used in summative evaluations [2]. Benefits of an iterative design and development approach are at hand; apart from the goal of developing excellent products that exactly fit the users' needs, the well-known saying "fail early, fail often" is a good synopsis of the fact that the cost of failure in an iterative design process increases as the project advances over time, while the risk of undiscovered issues declines the more iterations in the design and evaluation circle are carried out [8].

Unfortunately, the scientific vision of the ideal iterative development to support UX is rarely fully applied in industrial settings [11], [12]. On a general basis, UX is often mentioned as being important, but (in our daily practical work) already at the requirements phase details on how this "user experience" is to be built, supported or evaluated are already missing [5]. In the following we describe some of these insights from our daily work, to show shortcomings in the industrial practice when following iterative design and development phases.

2 An Experience Report from a Case Study from the field of IPTV

The main goal of the project used in this case study was to develop a user interface that supports UX in terms of positive emotions due to its seamless animations and quick feedback, that achieves high (user) ratings in terms of aesthetics and that users describe as stimulating and a system they would identify with. The goal was to develop a (set-top box) browser-compatible version of a user interface. Our intention was to deploy an iterative design and development approach including a set of evaluations. We describe in the following our experiences on why and how we failed to follow this approach.

2.1 Requirements Phase

What we learned during the project set-up phase and early requirement phase is that for supporting UX a simple textual description of requirements is not enough. We thus delivered a fully functional (flash-based) user interface prototype showing about 70% of the functionality, together with a requirements document of several hundred pages including task descriptions supported by user interface screenshots, detailed functional requirements as well as design specifications (colour codes etc.). The vision we had is the more details we provide, the better. Our project partners responsible for the implementation were simply overwhelmed, and given that they had a

functional prototype, preferred to use the prototype as reference (and simply did not follow the documentation in the beginning).

2.2 Design and Development Phase

Given that the design was already fully defined, we assumed that there was no further design phase necessary and went directly to the development phase. What became clear during the implementation phase with a partner company from Asia is that providing a full specification is overwhelming, while a slow development of the design and a simultaneous development might have helped to make our vision clearer.

2.2.1 Outsourcing

Outsourcing for the development sounds like a good idea, and sometimes it might even result in perfectly managed and performed projects, but there are also risks of spending more resources than expected. Various problems might come up if a project or parts of a project are outsourced.

Language

The first one is connected to communication and language problems when working with international teams or partners. While in general all involved persons can speak English as a common language, language skills and mastery vary, and there are chances that project goals (or design goals or rationales) are not understood correctly, or that misunderstandings occur when the responsible persons in the partner company forward information to their (internal) team.

What we found is that the internal communication in the development company was done in a foreign language, and back-translation to English changed the requirements and how they were interpreted.

Cultural Misunderstandings

Other issues that might occur are long feedback cycles if working with bigger international companies while one team at the same location might work more efficiently. Cultural differences in working attitudes and commitments might also be taken into account. For example, some cultures would tend to not say ‘no’ directly if something is not possible or feasible, while in other countries there might not be the same cultural understanding that a "no" is acceptable. In our case, we would have wanted the partner to show and communicate possible limitations, so we could change the design, but things that were

simply not implementable in the specified time frame were not named, until the very end of the project.

Acceptance Criteria

When outsourcing projects, we learned we would have benefited from writing down the criteria for acceptance in detail, as well as the consequences if one part of the acceptance criteria is not met or just partially met. This issue is also directly connected to the question how changing requirements are incorporated into the project and how these can be handled [6].

Overall, "committed" project management and communication on both sides is essential to transport goals and priorities to the respective teams. In terms of UX centred engineering, when outsourcing a project, it is advisable to have the project requirements ready early on (mind balancing over- and under-specification) and to reserve enough time for feedback cycles throughout the project to reduce the risk of misunderstandings and possible slower-than-expected progress, or if a project partner is not sticking to the agreed requirements.

2.2.2 Communication problems in multi-language and multi-location teams

In our modern world, it is not uncommon that the teams involved in a project are spread over various locations and have various mother tongues. Beyond that, sometimes team members also just need a person to ‘translate’ between their professional worlds, as e.g. the designer might have a different professional vocabulary than the software developer or the project manager, to avoid problems - like we experienced the software developer to wait for a rough framework for the app while the designer is lost in describing details that are at this stage unnecessary for the software developer. Other experiences we all might have made is the cliché of the software developer not focusing on usability (“it’s ‘perfectly understandable’”) or designers making beautiful yet unusable interface designs.

2.3 Evaluation Phase

Constant testing and user feedback throughout the whole product development process is desirable, but proved to be unrealistic. In reality, due to time or confidentiality constraints, UX engineering sometimes is more engineering to the CEO’s or client’s desires. The sooner the deadline for the final product, the more likely decisions might not be based on user studies and user evaluation, but will follow the decision of the client requesting the product, for example

the CEO of the client company deciding that the product will be pre-released at a major fair. These problems from time constraints (e.g. fair launch dates) lead to a change in planning where iterative development and evaluation are not scheduled accordingly any more, but the goal to "have something running" is more important than the usability or UX of the final system.

Due to the specific way we were involved in this project, it was not possible to have external people see the product (even if they would sign an NDA), so recruiting test participants was difficult and we ended up testing with people employed in the company.

3 Summary & Conclusion

This workshop paper gives some insight into our daily work and experiences we made during an IPTV-based software development that had as a main target to provide a new type of UX. The occurrences described where observed directly in the project we participated, which of course, we want to keep confidential in terms of participants involved. Although most of the topics presented in Section 2 are very subjective, we hope to provide some key lessons: (1) whatever the preparation, people need time to understand the scope of the work and to understand the level of quality required in the final product; (2) be aware of misunderstandings and differences in cultures - even if it seems obvious that the people speak the same language. What we learned as a central lesson is "to mind the gap" - between scientific lessons and industrial practice, between people's different roles and their understanding of the project - and also between languages - as sometimes a translation from French to English done by an Austrian, communicated by a Brazilian to an English speaking person from India, transcribing the requirements in a language from India, to be finally read by a French (native speaking) project officer will simply incorporate some surprises. Our very personal UX lesson learned: just smile :-).

References

[1] Bernhaupt, R. and Pirker, M. Evaluating User Experience for Interactive Television: Towards the Development of a Domain-Specific User Experience Questionnaire. In P. Kotzé, G. Marsden, G. Lindgaard, J. Wesson and M. A. Winckler (Eds.): Human-Computer Interaction – INTERACT 2013. 14th IFIP TC 13 International Conference, Cape Town, South Africa, September, Proceedings, Part II, pp. 642-659. Springer, Heidelberg. 2013.

[2] Dix, A., Finlay, J., Abowd, G. and Beale, R. Human-Computer Interaction (3rd Edition). Prentice Hall. 2003.

[3] Freebox gestural remote control (French site). Retrieved June 2015 from http://www.free.fr/adsl/freebox-revolution.html

[4] Kunert, T. User-centered interaction design patterns for interactive digital television applications. Springer Science & Business Media. 2009.

[5] Lanzilotti R., Costabile M. F. and Ardito C. Addressing Usability and UX in Call for Tender of IT Products. 15th IFIP TC.13 International Conference on Human-Computer Interaction - INTERACT 2015, Bamberg, Germany, September, pp. 14-18, Springer, Heidelberg. 2015.

[6] Lieberman, H., Paternò, F., Klann, M. and Wulf, V. End-user development: An emerging paradigm, pp. 1-8, Springer Netherlands. 2006.

[7] Löwgren, J. Applying design methodology to software development. In Proceedings of the 1st conference on Designing interactive systems: processes, practices, methods, & techniques - DIS '95, Gary M. Olson and Sue Schuon (Eds.). ACM, New York, NY, USA, pp. 87-95. 1995.

[8] Mayhew, D. J. Principles and Guidelines in Software User Interface Design, Prentice-Hall, 1992.

[9] Orange LivePlay gestural remote control (French site). Retrieved June 2015 from https://assistance.orange.fr/tv-decodeur/tous-les-decodeurs/installer-et-utiliser/associer-la-telecommande/telecommande-noire-tester-la-commande-gestuelle_24889-25466

[10] SFR, La Box Fibre (French site). Retrieved June, 2015 from http://www.sfr.fr/la-box-tres-haut-debit/#sfrintid=V_nav_box_la-box&sfrclicid=V_nav_box_la-box

[11] Venturi G. and Troost, J. Survey on the UCD integration in the industry. In Proceedings of the third Nordic conference on Human-computer interaction - NordiCHI '04. ACM, New York, NY, USA, pp. 449-452. 2004.

[12] Vredenburg, K., Mao, J. Y., Smith P. W. and Carey T. A survey of user-centered design practice. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '02. ACM, New York, NY, USA, pp. 471-478. 2002

[13] Wright, P., Blythe, M. and McCarthy, J. User experience and the idea of design in HCI. In Interactive Systems. Design, Specification, and Verification, pp. 1-14. Springer Berlin Heidelberg. 2006.

A Review of Milestones in the History of GUI Prototyping Tools

Thiago R. Silva, Jean-Luc Hak, Marco Winckler
ICS-IRIT, Université Paul Sabatier, Toulouse, France
{rocha,jean-luc.hak,winckler}@irit.fr

Abstract. Prototyping is one of the core activities of User-Centered Design (UCD) processes and an integral component of Human-Computer Interaction (HCI) research. Nonetheless, for many years, prototyping was synonym of paper-based mock-ups and only more recently we can say that dedicated tools for supporting prototyping activities really reach the market. In this paper we propose to analyze the evolution of prototyping tools proposed by the academic community to support research activities and prototyping tools that are aimed and improve the development process of commercial user interfaces. Thus, this paper presents a review of past and current graphical user interface prototyping tools, in order to set up the state of the art in this field, observing fundamental milestones of features over time. For that, we have screened publications presented since 1988 in some of the main HCI conferences and 113 commercial tools available on the web. The results enable a brief comparison of characteristics present in both academic and commercial tools, how they have evolved and what are the gaps that can provide insights for future research and development.

1 Introduction

Prototyping graphical user interfaces (GUIs) is considered one of the most important activities in a User-Centered Design process (UCD) as a mean to investigate alternative design solutions in the early and advanced phases of development. For this purpose, a lot of tools have been developed to support a set of activities, such as planning, sketching, designing and evaluating prototypes of user interfaces.

Prototypes, particularly in the early stage of designing GUIs, are an important way to communicate and discuss ideas and requirements [12]. Low-fidelity techniques are often used by designers to sketch and present new ideas and concepts about the interface that will be built. This kind of activity involves the users early in the design process, promoting an effective participatory design and improving the user experience. With the advance of software development lifecycle, medium and high-fidelity prototypes can be used to refine some features or provide more accurate information about interaction options.

In the academic context, since 1988, the most important conferences in the field of Human-Computer Interaction (HCI) have given space for tools

developed in order to solve several scientific challenges related with this theme. However, dedicated tools for supporting prototyping activities only started to have an impact in the market by 2003. Thus, we can observe a temporal gap between the research interest and market adoption.

The aim of this work is to investigate the state of the art in GUI prototyping tools, analyzing the main contributions in terms of new ideas and features, regarding the main milestones over time, as well as identifying new research gaps in this area. To accomplish that, we described a review of tools that have been presented to the academic community or provided in the commercial context. For academic tools, we systematically investigated the main conferences on HCI since 1982, with the first SIGCHI Conference. For commercial tools, we investigated the most common ones, constantly mentioned by designers and present in most professional activities related to prototyping. The next section of this paper presents the research protocol used to investigate tools in both academic and commercial context.

2 Methodological Approach

In this section we present the methods used for the analysis of the academic and commercial tools.

2.1 For Academic Tools

We sought top ranked HCI conferences and selected those that were sponsored or co-sponsored by ACM, IEEE and/or IFIP. Only proceedings published in English and available in digital form were considered. Domain-oriented conferences such as mobile, embedded, robot, pervasive and ubiquitous interfaces, were excluded from the analysis. We considered the following conferences and period of publication: ACM CHI (1982-2014), ACM UIST (1988-2014), ACM DIS (1995-2014), ACM EICS (2009-2014), IFIP INTERACT (1984-2013).

We screened papers that propose prototyping tools and/or describe the use of existing tools for supporting prototyping. We considered relevant those papers that included in the title and/or abstract any of the following keywords: prototype, prototyping tool, prototyping interface, wireframe, wire-framing, sketch, sketching, draws and drawing. With these keywords, 7.243 publications were selected. Subsequently, we excluded papers reporting tools developed for specific prototyping in specific environments (ex. sketches of buildings for architects, drawings for designers, circuits and physical devices for engineers and so on). We also did not consider papers that report model-

based prototyping of multimodal user interfaces because our main interest is in tools that can support the concrete development of user interfaces, not only to model it. Finally, we mainly considered full papers. However, one tool that was not published as a full paper, but describes important and concrete results for this field, was added to the final list of papers.

2.2 For Commercial Tools

There are a large number of commercial tools to prototype GUIs. 113 prototyping tools have been examined in two steps. The first step was to check the main features of each tool and roughly categorize them. Doing so, we checked the website of each tool, sorting them according to their similarities with other tools (e.g. the way they handle the creation of the user interface, the precision that can be achieved when describing the behavior of the prototype) and removing from the list each tool that does not match with the study. Then, we have defined some criteria that tools must comply with before doing further analysis. The criteria were: “Is the tool a standalone software or an extension/library/framework?”, “Is prototyping generic interfaces possible?”, “Is there a free trial of the tool?”, “Is the tool still updated and documented?” and “Does the prototype produced with the tool support any interaction?” A few exceptions were made if the prototyping tool was featuring interesting functionalities, but eventually did not accomplish one or more criteria. In a second step we focused on the mechanisms used by prototyping tools for the creation of a prototype and especially the construction of the presentation and the dialogue.

With those criteria, we discarded 23 tools that did not draw attention during the first analysis. Some of these tools were not exactly a software tool, but just a library or theme to apply on existing software, so we focused on this existing software and its features instead. We also removed from the list tools that cannot produce a reusable file. The advantage in using software to produce a prototype is that they can easily be modified, reused and shared. Some tools also were not still updated and documented, so they were eliminated; as well as tools that were too specific and too restrictive, so prototyping generic interfaces was not possible.

3 Selected Tools

We have selected 17 papers really talking about academic prototyping tools which included the following tools: SILK [5], DENIM [6], DEMAIS [1] and CogTool [4] (from ACM CHI), Gambit [11] (at ACM DIS), GRIP-it (at ACM

EICS), Mirage [7], Ensemble, Lapidary [8], Druid and Monet (at ACM UIST), SIRIUS, MoDE, SCENARIOO, Freeform and SketchiXML [2] (at IFIP INTERACT), and ActiveStory Enhanced [3] (from XP International Conference). We have retained 90 commercial tools for detailed analysis. We have identified three categories depending on what can be prototyped with the tool: the behavior, the presentation (visual aspect) or both.

The first one gathers 9 tools that are more suited for representing the behavior of a prototype. In the second one, we have regrouped 8 drawing tools like Inkscape or Photoshop, where it is possible to create a visual prototype without caring about the behavior or the interactions possible. Finally, the last category corresponds to tools that can manage both graphical and behavior aspects and it features the remaining 73 tools. Mock-up tools and wireframe tools fall into this category. Therefore, we have decided to focus on this last category since they are mainly tools that are dedicated to the construction of fully functional prototypes.

4 Set of Milestones Observed

A temporal view of academic and commercial tools is presented in Figure 1. We can observe three main periods of interest in this kind of tool.

The first one, before 1995, coincides with the emergence of UIMS tools. The first tools mainly treated of high-fidelity prototypes, using mostly design elements from the final interface, and being strongly dependent on the hardware. The main advantage of the UIMS tools is that, after development and testing, the interface prototype can be directly attached to the application the prototype becomes the interface [7]. Since the emergence of UIMS tools, authors start to discuss the lifecycle of development processes using prototyping tools, redesign is discouraged because it takes time and is perceived as a “waste" of money, and it is suggested that evaluation should occur in early phases in the development process. Nonetheless, UIMS tools do not give the flexibility needed in the early stages of prototyping. Typically, in the brainstorm meetings, designers are more focused on describing important aspects of the problem to be solved, more related to requirements than to interface aspects. Even if one can rapidly build a prototype to demonstrate his/her ideas using this kind of tools; they won’t be able to deliver a solution that does not consider design aspects which are unnecessary for the early phases. Tools such as PowerPoint and Visio are often reported as a mean to support the edition of the presentation of user interfaces; even though these tools should be generic.

The second period (1995-2005) we can observe in Figure 1 is represented by tools in which one can really draw a functional prototype; some of them support sketching. SILK [5] and DENIM [6] are the first tools designed for hand sketched GUIs and are still the most representative tools that use sketch recognition. Landay et al. [5, 6] believe sketches are important for communicating ideas with other team members and gaining valuable feedback early in the design process.

This period was followed by an increasing interest in other ways to prototype interfaces and the inclusion of behavior modeling which can be roughly dated after 2005, leading up to a third and last moment, with a substantial increase of commercial tools from 2007 to now.

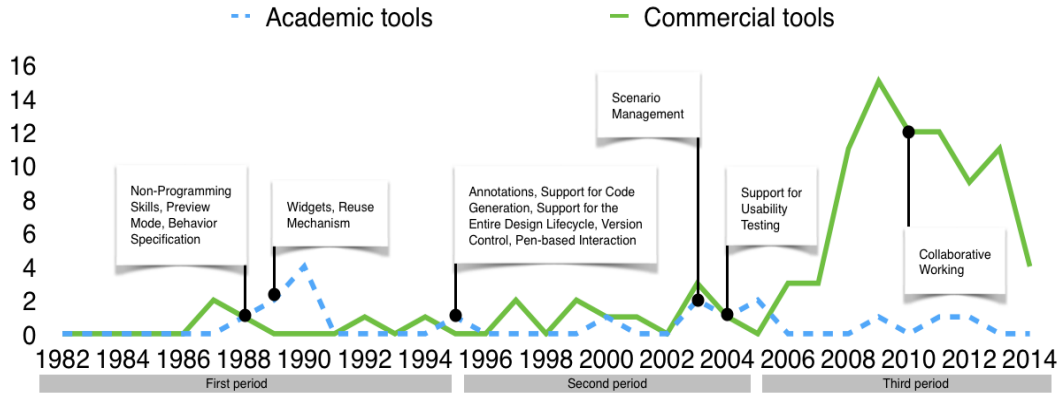


Figure 1. Number of both academic and commercial tools per year.

Hereafter, we describe some of the main features, set in the time by milestones, observed during the analysis of both academic and commercial tools.

4.1 Non-Programming Skills

UIMS tools started providing non-programming skills with resources that allow designers to build their interfaces without programming skills. The goal was to allow users to concentrate on what is to be done rather than how to do it [7]. One way to accomplish this objective is to give users the ability to directly manipulate the representations of concepts from the task domain (e.g., design objects). Furthermore, authors of UIMS tools believe that special purpose systems are more likely to provide the kind of semantic and articulatory directness necessary to allow the user to concentrate on the task rather than the tool. Examples of tools pursuing this goal include MIRAGE [7] using objects, and its successors Lapidary [8], Ensemble and Druid. Most of

the tools that came after this milestone do not require programming abilities from users, even with Lapidary [8], for instance, demanding some Lisp programming ability to express more refined behavior. Nowadays, it is common sense between developers of GUI tools that they should simplify the activity of designers and interface engineers, and requiring some level of programming skills is a throwback. Because of that, all the tools analyzed work with abstract elements and behavior models as a way to provide prototyping resources for the users, without requiring any kind of ability to program software. We believe that this is a well-established feature today.

4.2 Pen-Based Interaction

Pen-based interaction regards the possibility to interact with prototypes through a pencil and paper metaphor. This kind of interaction uses an electronic pen to really draw the interface in a touch device. Landay and Myers [5] believe that designers feels more comfortable using pencil and paper to prototype initial versions of their ideas, discuss and present them in brainstorm meetings.

The first time this feature appeared was in SILK [5], followed by DENIM [6], Freeform, SketchiXML [2], ActiveStory Enhanced [3] and Gambit [11]. Authors of these tools observed that all designers sketch with pen on paper as a regular part of their design process, even though eventually all of them end up using computerized tools. This activity allows designers to explore the space of possibilities more effectively through sketching than using computer-based tools, at least during the early parts of the process.

Given sketching characteristics, normally prototypes of this kind are throwaway. Additionally, features using sketch recognition are also provided, in a way that sketches made by hands can be recognized and interpreted by the tool. In this case, widgets (in a higher level of fidelity) are automatically generated to support further evolutions in the prototype that might not be throwaway. Such kind of technique is presented in all of the tools listed above.

However, even more than 20 years after SILK [5], there are no commercial tools that implement this kind of feature using sketch recognition. Tools like Blueprint, Cacoo, Mockup Plus, NinjaMock and Pidoco support both palette and sketching methods of interaction, but not sketch recognition. Even so, it is definitely an unpopular feature among commercial tools.

4.3 Widgets

Widgets have been used to build prototypes since Lapidary [8]. Their use guides the major part of tools that work with a palette as interaction technique

nowadays. Widgets have the advantage to facilitate the process of element manipulation, offering a fast manner to set various components as menu bars, buttons, windows and form elements. Even tools that work with a sketching mechanism like SILK [5] and DENIM [6] set up a library with known elements (drawn before) and treat them as a widget for future uses. The restriction of a palette of widgets is that the prototype is limited to the components that are available on the prototyping tool and their representation. For instance, Balsamiq provides only low-fidelity widgets.

4.4 Behavior Specification

Prototyping tools facilitate the representation or the implementation of interactions and actions. This behavior specification makes it possible to define higher-precision and evolutionary prototypes, though the tools are not equal in the possibilities they offer.

As we have only analyzed academic tools with clear purposes of prototyping, almost all of them provide some kind of behavior specification. Lapidary [8] was the first one we noticed. It provides interesting features and resources that lead to dialogue construction. The dialogue can ensure the task sequencing and so the different screens of the interface. Other behavior building mechanisms using event handling on widgets can set up more refined behavior, too. Animations are another way to define elementary behaviors through a predetermined scenario.

It is possible to distinguish tools by the way they represent the dialogue. Tools like Pidoco represent the sequencing of screens through a state machine or a diagram. Tools can use similar mechanisms to ActiveStory Enhanced [3] or Balsamiq which support only basic interactions with hyperlinks between screens. Tools like Axure or JustInMind can use conditions, variables to modify properties and states on top of the hyperlinks. Finally, some tools like Appery.io and ScreenArchitect support the use of programming code.

4.5 Collaborative Working

In our study, we have found 25 tools that feature functionalities allowing designers to collaborate on the same prototype. Appery.io and Hotgloo are among the first prototyping tools that support collaborative work. Collaboration is made possible in two different ways: in a synchronized environment and in an asynchronous environment. Tools that support synchronized modifications like HotGloo or InVision allow several users to work at the same time on the same prototype. Every modification is applied on each instance of the prototype. On top of that, functionalities that help users to

know what collaborators are doing are often available like a tele-pointer or a chatroom. Tools that support asynchronous collaboration like Axure or GUI Design Studio are mainly based on revision control where each user has a local copy of the prototype and annotations.

As for academic tools, Gambit [11] is really dedicated to collaborative working and supports sketch production and visualization on different devices, session storage and retrieval, private and/or public production of sketches, over providing a broad view of the drawings (like papers arranged on a wall) and a fine view of them.

4.6 Reuse Mechanism

Reuse is always important to reduce the workload of designers and users. In the prototyping domain, tools generally use widget libraries to provide reusability: they provide templates and pre-defined behaviors. Sketching tools like SILK [5] and SketchiXML [2], by the use of UI widget representations, support the reuse mechanism of interface elements which have already been taught to the system before. DENIM [6] and Gambit [11], in contrast, do not support reusable components.

Commercial tools like Appery.io, HotGloo, iRise, Protoshare and UXPin feature the usage of breakpoints and screen versions, thus promoting reuse of design for multiple devices. This method supports multiplatform design by resizing and repositioning elements that have already been created. It is also possible for some tools to apply a theme on the prototype like ForeUI or MockupScreens. Doing so will switch the appearance of the prototype without recreating it. Another way to reuse existing work is to import an image of a prototype designed using any drawing tool like a paper prototype. After that, the prototyping tool is used to add interactions for the imported prototype. Some commercial tools are based on this mechanism of import and interaction adding like MarvelApp, Flinto or Notism.

4.7 Scenario Management

Scenario-based design is a family of techniques in which the use of a future system is concretely described at an early point in the development process. Narrative descriptions of envisioned usage are then employed to guide the development of the system and the conduct of experiences [10].

Scenario management refers to the ability of the tools to work with different scenarios and manage them in an integrated way with prototypes and behavior descriptions. It is not an easy feature to implement because it is strongly dependent on the complete development processes and their models, so their

implementation normally becomes too restrictive. Despite the fact that this feature has appeared first in Freeform, a restrictive tool working as a Visual Basic 6 plug-in, in 2003, there has not been much evolution since then. If scenarios are seen as simple annotations in the prototype, we can consider several tools providing some kind of feature to treat them, but if we see those scenarios like complete requirements specifications, which need to be managed and controlled through the whole lifecycle, there are currently no tools able to solve this problem.

4.8 Preview Mode

Preview mode is an important feature to enable visualization of one executable version of the prototype. In this mode, we can execute and simulate all interactions specified during the construction of the prototype. Users can test the application as a rough final product. It is important, in this case, to visualize how the prototype will really appear in a real environment and to promote usability testing and collect adequate feedback from particular stakeholders. MIRAGE [7], Lapidary [8], SILK [5] and DENIM [6] feature this mode whereas a third application is needed for SketchiXML [2]. Most commercial tools include a preview mode.

4.9 Support for Usability Testing

Prototypes can be used to support usability tests by collecting data from users. Indeed, it is possible to store useful information that can be measured while the prototype is being used (time spent on each screen, the area clicked, etc.). Tools like CogTool [4], Solidify, PickFu, IntuitionHQ provide functionalities that help to create usability tests. For instance, it is possible to add instructions or questions to the test of the prototype and to create tasks that have to be accomplished. With ActiveStory Enhanced designers can export the prototype to a web-based Wizard-of-Oz testing tool, allowing test participants to remotely walk through a UI while recording metrics such as mouse movements and time spent on pages. After carrying out user tests, collected data are made available with an interface dedicated to the management of results.

4.10 Support for Code Generation

Another advantage of prototyping using a software is that it is possible to automatically generate code that can be used directly or refactored afterwards. SILK [5] generates code for an old OpenLook Operating System and Freeform

for Visual Basic 6. Both of them are very restrictive in terms of environment. SketchiXML [2] and Gambit [11] produce interface specifications and generate code in UsiXML, an open-source format based on XML.

Among all commercial tools that have been studied, 21 of them can generate, for example, web pages based on the prototype created. Well-known tools like AppSketcher, Axure, ForeUI and JustInMind directly generate these pages, considering they use code for dialogue descriptions, in this way providing a mechanism to support other phases of the design lifecycle and evolving the prototype to the final user interface.

4.11 Version Control

Version control is often related to the fact that any document or software that is created can be modified at several times during its lifecycle. It allows each user to check the current state of the document, different versions that exist and the reason of any modification. Version control is a mechanism that is interesting for prototyping since the prototype is constantly evolving due to feedback, needs and requirements that emerge throughout any project, and because a prototype can be declined in several versions depending on the options of design that are considered by the designers.

SILK [5] supports version control through design history, and started using this feature in 1995. DENIM [6], in contrast, does not support version control. Another interesting feature when designing several solutions and options of design is the ability to compare two versions. Commercial tools like Alouka, Codiqa, FluidUI, HotGloo and JustInMind support only version control. Concept.Ly is able to compare two different screens using a slider.

4.12 Annotations

The annotation system is an interesting feature since it may be a way to collect various feedback on problems that are identified by the annotators and to communicate with users [9].

Annotations are dated information and can be used to keep a trace of the evolution of a prototype since they might influence the development. That information can take several forms like inquiries, decisions, constraints and specifications, use cases of a software, problems encountered by users, advantages found in other software, test data or even ideas.

Several ways of annotating have been implemented in prototyping tools. For instance, SILK [5] and DEMAIS [1] supports textual annotations as an input design vocabulary. Alouka, Balsamiq, inPreso, Lumzy and WireFrame Sketcher support annotations through widgets, the simplest method. Axure,

MockupScreens and JustInMind support this feature as a property of widgets. There are also tools that have a dedicated annotation mode like Concept.Ly, ForeUI and NinjaMock.

4.13 Support for the Entire Design Lifecycle

Support for the entire design lifecycle means that designers can work with the same prototype since the early stages of development, evolving it towards most refined levels and becoming the final user interface through an evolutionary and iterative development lifecycle. This characteristic is a frequent constraint to adopt current prototyping tools in several development processes. The time wasted building a throwaway low-fidelity prototype becomes an adverse argument to obtain stakeholders support.

SILK [5] supports the transformation process of the sketches to real widgets and graphical objects, but no more than that. SketchiXML [2] and Gambit [11], even being a sketching platform, need to be integrated with other UsiXML tools to support several levels of prototyping. Thus they need a third application to provide that. DENIM [6] and DEMAIS [1] do not support different refinement levels, so they do not cover the whole lifecycle (they do not generate finished HTML pages, for example). DENIM [6] just allows the navigation among different representations in a web-design prototype, such as site maps, storyboards and mock-ups.

Some tools like ScreenArchitect support model description by providing links between prototypes and models like state machines, leading then to a more integrated environment in the UCD development processes.

5 Conclusions

We can observe these milestones as a large spectrum of features being covered by prototyping tools over the time. As we have supposed, original and innovative features come from academic tools, generally providing solutions to problems that will be addressed by commercial tools some years after. We have also observed, in particular cases, that some features like pen-based interaction, presented by SILK [5] twenty years ago, seem to be not interesting to be adopted by tools used in the commercial context.

Another aspect we can highlight is the pool of commercial tools launched after 2008, when Balsamiq came up. These tools have incorporated the most aspects we report in this paper, providing, in different levels, implementations of these concepts, and many times, being strongly repetitive in their qualities. Nevertheless, it shows a continued interest from both the academic and

industrial community in this theme, suggesting an open space of research in several points.

This is a research still under development and is part of an initial investigation about prototyping tools. This state of the art on existing prototyping tools will help us to have a better understanding of the remaining gaps of features that can support the software development process through the whole lifecycle of prototyping.

Features and current directions point to an accurate analysis of the main gaps of features and open research problems, based on prototyping as support activity for the development lifecycle. Regarding these gaps, we have already identified little support of the tools for annotation activities in a requirements process. Tools that treat annotations as a property and not as a single remark put in the prototype better support the specification process of requirements. Even though, the way they capture the information coming from this annotations is not profitable to be used for supporting business rules, specification of needs or functional descriptions.

Another important gap already identified is related to integrated support for development models. Task and system models, when considered, are normally no integrated to the prototyping activity in most of the current tools. At this way, it is hard to work in an integrated environment where it is possible to build low-fidelity prototypes, evolve them to more refined ones and, from scenarios, requirements annotations and constraints, support the development of models and check the user interface according to a unified prototyping specification.

References

[1] Bailey, B. P., & Konstan, J. A. (2003, April). Are informal tools better?: comparing DEMAIS, pencil and paper, and authorware for early multimedia design. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 313-320). ACM.

[2] Coyette, A., & Vanderdonckt, J. (2005). A sketching tool for designing anyuser, anyplatform, anywhere user interfaces. In Human-Computer Interaction-INTERACT 2005 (pp. 550-564). Springer Berlin Heidelberg.

[3] Hosseini-Khayat, A., Ghanam, Y., Park, S., & Maurer, F. (2009, January). ActiveStory Enhanced: Low-Fidelity Prototyping and Wizard of Oz Usability Testing Tool. In *XP* (pp. 257-258).

[4] John, B. E., Prevas, K., Salvucci, D. D., & Koedinger, K. (2004, April). Predictive human performance modeling made easy. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 455-462). ACM.

[5] Landay, James A., and Brad A. Myers. "Interactive sketching for the early stages of user interface design." Proceedings of the SIGCHI conference on Human factors in computing systems. ACM Press/Addison-Wesley Publishing Co., 1995.
[6] Lin, J., Newman, M. W., Hong, J. I., & Landay, J. A. (2000, April). DENIM: finding a tighter fit between tools and practice for Web site design. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems (pp. 510-517). ACM.
[7] McDonald, James E., Paul DJ Vandenberg, and Melissa J. Smartt. "The mirage rapid interface prototyping system." Proceedings of the 1st annual ACM SIGGRAPH symposium on User Interface Software. ACM, 1988.
[8] Myers, B. A., Zanden, B. V., & Dannenberg, R. B. (1989, November). Creating graphical interactive application objects by demonstration. In *Proceedings of the 2nd annual ACM SIGGRAPH symposium on User interface software and technology* (pp. 95-104). ACM.
[9] Naghsh, A. M., Dearden, A., & Özcan, M. B. (2006). Investigating annotation in electronic paper-prototypes. In Interactive Systems. Design, Specification, and Verification (pp. 90-101). Springer Berlin Heidelberg.
[10] Rosson, Mary Beth, and John Millar Carroll. "Usability engineering: scenario-based development of human-computer interaction". Morgan Kaufmann, 2002.
[11] Sangiorgi, U. B., Beuvens, F., & Vanderdonckt, J. (2012, June). User interface design by collaborative sketching. In Proceedings of the Designing Interactive Systems Conference (pp. 378-387). ACM.
[12] Schvaneveldt, R, W., et al. "Towards a Modular User Interface" (CRL Technical Report No. MCCS-85-10). Computing Research Laboratory, New Mexico State University, Las Cruces, New Mexico. 1985.

Workshop:
IFIP WG 13.7 Workshop on Designing Interaction and Visualization for Mobile Applications (DIViM 2015)

The Future of Visual Perception

Gerrit C. van der Veer
Sino European Usability Center
Dalian Maritime University
Dalian, China
gerrit@acm.org

Abstract. We sketch some Psychological and Anthropological notions on visual perception, ranging from automatic mechanisms, to learning and context. Moreover, cultural knowledge supports context-specific understanding. New technical developments trigger changes in how and what people may perceive as well as what they let others perceive. This requires reconsidering privacy. And interactive VR worlds still await facilities for multi-person interaction.

1 Introduction

Daily use of information and communication technology (ICT) by non-experts in computer science is quickly spreading in our world. Literally billions of people [5, 7] communicate, work, travel, drive cars and use household appliances, use facilities for health care and learning, participate in culture and social life, and make use of their civil rights, by looking at, and interacting with, screens of any size. Rendering information and affordances for interaction on screens requires understanding of how people look, understand, and be guided in their interaction, by what is displayed.

This paper is an analysis of how people perceive visual images, how perceptions is influenced by context and by learning, and how people's understanding of what is being rendered has been changing over the centuries and is rapidly changing in the current era of revolutionary ICT innovations and applications. We will illustrate that cultures with different values or different views on aesthetics require different ways of representation.

In the current turmoil of human mobility and geographical cultures interconnecting, of social media being deployed as soon as replaced by new hypes, cultures develop and change rapidly. Designers of hardware, software, and applications for users need to be aware of these mechanisms and developments in order to optimally serve the users.

In 2 we will give some highlights regarding what occurs before stimuli of the senses results in perception. In 3 we show how basic processes as well as learning affect attention to sensory input and the resulting meaning. Cultural influences to this process are illustrated in 4, and in 5 we show how new techniques change perception and interaction with visual rendering.

2 Before Perception

Prior to any perception, humans receive input through the senses, of which sight is only one (among the others most scholars count at least: hearing, taste, smell, and a complex of sub-senses like touch, thermoception, nociception, kinaesthesia, proprioception, and equilibrioception). Senses are the physical processes of stimuli (light, sound, etc.). For the sense of sight, the stimulus is light, received by several types of photoreceptors in the retina [6]. The image received differs between the two eyes, hence the signals transferred to the brain allow for 3D vision. The image remains available in the receiving organ for a short period (less than 1000 milliseconds). This time cannot be prolonged by rehearsal, since attention does not affect the physical process. All sensors only act if the stimulus is above a certain lower limit (the absolute threshold) and with increasing intensity the receptor reacts stronger until an upper limit is reached. Above this limit stimulation will result in a possibility of damaging the sensor and an experience of pain (this is why people quickly learn not to look directly to the sun). In order for two stimuli to trigger a distinguishable reception, the difference needs to be beyond the "just noticeable difference" threshold [11]. If a sensor is exposed to a certain stimulus for a long time, it adapts and gets less sensitive. On the other hand, if the stimulus changes suddenly the sensor reacts quickly.

A place where theory about senses has been put into practice is the hospital. The colours in the operating room (OR) are usually greenish. The reason for this is the so called afterimage effect of the eye. When the eye has been exposed to a lot of red colours (like blood), it will start to adapt to the colour red (get less sensitive to it). If you look in a white area directly after being exposed to intense red colours, your image will be greenish. This is because your eye still is less sensitive to red, so the image lacks red colours to compensate the green. Doctors won't notice this greenish afterimage, because OR's are coloured green. The afterimage then simply blends with the environment, and is less annoying.

3 Perception, Learning and Context

Perception is the result of attention, and only what is processed by the senses can subsequently be perceived. Schacter et al [13] define perception as "the organization, identification, and interpretation of sensory information in order to represent and understand the environment". Gestalt psychology is a theory about fundamental laws of human perception, explaining the human ability

automatically structure stimuli into meaningful perception [15]. Some important Gestalt laws in the domain of vision are:

- Closure - we tend to see complete figures, even if some parts are missing. Like the square on the right: it is merely four circles with a gap in them, but we perceive it as square (see Figure 1a);
- Similarity - when we see similar objects (colour, texture, shape, size, orientation) we tend to group these objects together. So in the image on the right we tend to see diagonal stripes (see Figure 1b);
- Continuation - if two lines cross each other, we automatically assume that each line follows it initial path, because that would be a logical continuation of the lines (see Figure 1c).
- Proximity - depending on the proximity of objects that are ordered in rows and columns, we will either perceive them as a whole, as a bunch of rows, or as a bunch of columns. Figure 1d is perceived as a whole; Figure 1e as 'rows', and Figure 1f as 'columns'.

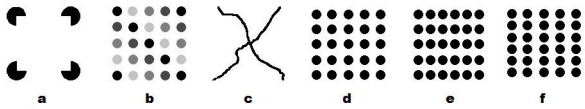


Figure 1. Illustration of Gestalt laws Closure (a), Similarity (b), Continuation (c), and Proximity.

In general this type of phenomena supports interpretation. On the other hand there are visual illusions [2] that might sometimes disturb perception: the horizontal lines at the left and the circle at the right of figure 2 cannot be perceived as they are in fact drawn.

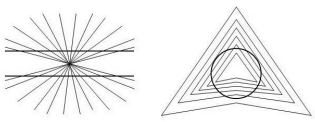


Figure 2. Some visual illusions.

The above mentioned types of phenomena will be found for people from many cultures, independent of education or explicit learning. On the other hand, perception of renderings that are not perfectly clear or precise is often supported only for people who have relevant basic knowledge (e.g., language, Figure 3-left).

Figure 3. Ambiguous text (left); Effect of colour contrast (middle); Spoiling capitals (right).

When designing for a screen, obvious knowledge of human perception should be applied, including the effects of contrast on colour (Figure 3-middle), and the meaningless use of capitals on readability, as illustrated in Figure 3-right.

And be aware, readers from screens follow the "normal" reading behaviour of their language culture: in the Western world from left to right, top to bottom, but often scan only the left few words or symbols and never see lower right info.

4 Cultural Knowledge Makes Us Understand

If we understand the logo in figure 4, this is because we are able to read the characters ("solidarnosc") and associate this with the English "solidarity", have seen enough pictures of political demonstrations to interpret the typography, and remember (or have read about) the Polish shipyard strike in 1980 and the following political change in 1989.

Figure 4. Logo of Solidarnosc, the Polish revolt 1980-1989.

4.1 Perspective

In Western society, for about 500 years (until Cubism challenged the relevance of it), perspective drawing has been the standard base of rendering visual reality on a 2D surface. Alberti and Brunelleschi are generally considered the original developers of the technique, around 1430.

Figure 5. New perspective, Florence 1430 (left); and isometric rendering, London 1616 (right).

Figure 5-left shows a painting by Masaccio in the Brancacci Chapel, Florence, ca. 1430 [14]. The new way of rendering was immediately

considered to provide more realistic representations. However, even when the technique was known and taught, there could be conditions when a deviation towards isometric representation of reality made much more sense if details really mattered for the viewer, see Figure 5-right [9], where even today specialists in architecture can derive details and measures of St. Pauls in 1616 that would not be found even in a photograph.

In the Western world Chinese visual art could be misunderstood as being primitive on perspective. However, the techniques and rules are well documented and taught, and have been preferred by the culture for many centuries [10]. In addition, when we consider the Chinese genre of scroll painting, this way of rendering makes in fact much more sense for whoever is viewing (while walking along) a horizontal scroll, see Figure 6.

Figure 6. Spring Morning in the Han Palace – part of the scroll, and a fragment showing the Chinese perspective [12].

4.2 Humans and Gods

Figure 7. Pietro Lorenzetti, Compiante (detail), [1].

In the European Middle Ages rendering pictures of humans included representing their social or religious role or state, by showing indicating attributes, dress, etc., see Lorenzetti's Compianto in Basilica Inferiore, Assisi (about 1320, Figure 7) where all saints are characterised by a Halo and the Virgin Mary is further specifically indicated by the colour of her dress.

4.3 Aesthetics

One important aspect of the cultural background of perception indicated above is aesthetic preferences. To a certain extend these just result in positive or negative feelings towards the rendering, which might or might not strongly affect the meaning of the perception [3, 4]. Early Egyptian drawings and paintings show remarkable standard ways of representing humans and gods (and combinations of these) [8]: female Pharaohs feature a beard, faces are shown strictly in profile, even if the crowning set of bull horns (indicating Apis) is always drawn from the front. It seems that in this case appropriateness of representation also supported easy recognition of the intended meaning (human or animal face, deity).

5 Modern Techniques Trigger Novel Perceptions

Presentation techniques evolve fast. We find developments in rendering that support 3D perception; perception of history; unobtrusively collecting knowledge of invisible characteristics; and perception of personal and private knowledge about strangers.

5.1 3D Perception – to be Continued

The 19th Century showed the development of panoramic (still) paintings that provided immersive experience of being in a 3D world. One century later panoramic movie projection upgraded this to a dynamic experience. Early 3D vision required special viewers, like in Figure 8, or polarized presentation with matching goggles. Currently, 3D vision can be provided in several ways, e.g., with goggles, with the help of polarised glasses, or by holograms. Newer technology allows "normal" vision where the presenting screen carefully aims different images at both separate eyes. Developments are still going on to get rid of those prostheses that detach the viewer from the real visual world around.

Figure 8. 20th Century 3D picture requiring red-green glasses (left) and a separation viewer (middle) for two separated 2D images (right).

5.2 Looking into Past and Future

Who is interest in the past of an object or in scenes related to cultural or natural history will increasingly find possibilities to watch renderings of what might have been a state at a specified number of years ago, by looking at a specific App on a smartphone or an interactive screen based museum device. Advanced applications allow the viewer to specify the time and aim the phone or screen at a scenery or artefact, seeing the presumed past state in its (current) context. IKEA applies Metaio, an App to view how new pieces of furniture would fit in people's living rooms. For private use Google's Cardboard box allows VR presentations to run on a smartphone.

5.3 Looking What's Inside

The British Museum provides a non-destructive opportunity for visitors to find out what is inside the mummy, allowing to virtually unwrap the artefact through layers of X-ray and CT representations. Medical professionals use these techniques for diagnostic, radiation, and surgical purposes, and these applications are used by non-professionals as well: Apps allow future parents to watch their unborn child on their smartphone, and many of these videos are shared on YouTube, even if the interpretation is not straightforward without careful explanations of an expert.

5.4 Looking to Create New Knowledge for the Benefit of Others

Supermarkets are currently experimenting with eye-tracking glasses fitted with a camera (e.g., to be obtained from Amazon), to study looking behaviour of customers. So far the customers supposedly are aware of the destination and use of the knowledge they provide, but privacy laws lack behind the new possibilities.

5.5 Looking may Turn into Spying on Privacy

Google glass, now (temporarily?) withdrawn from the consumer market, came with experimental Apps that allowed, for example, identification of faces based on facial characteristics stored in the cloud. The wearer of the Glass could be informed of the identity and other information of the, until then unknown, stranger. And we all know how hard it is to get personal information removed from data sets collected and stored in the cloud. Here, again, laws lack behind possibilities. But experience on new applications (e.g.,

of social media) show users quickly get used to the fact that their information moves around beyond control, and stop bothering about this till serious damage actually occurs to their privacy.

6 Designing for Future Perception: the Good, the Bad, and the Ugly

Designers should be aware of effects and mechanisms indicated above. In a world where people interact, perceive, and communicate with multiple screens and mobile devices, visual renderings should fit psychological processes as well as varied and rapidly changing cultures and context of use.

The good news is that new applications emerge rapidly, VR and Augmented Reality can be offered on location and to go on an increasingly diverse type of platforms.

The bad news concerns the indicated aspects of privacy that still needs to be considered and will require reconsidering national as well as international legal measures and a new professional orientation on ethics of design and application of novel technology.

Ugly is still the lack of technical facilities for sharing a VR world. Even the lightest of VR glasses like Oculus Rift still detaches the viewer completely from the social surroundings, and collaboration in VR in fact would require the use of multiple CAVEs. Technology experts should find their challenge here.

References

[1] commons.wikimedia.org/wiki/File:Pietro_lorenzetti,_compianto_(dettaglio)_basilica_inferiore_di_assisi_(1310-1329).jpg. Downloaded June 9, 2015

[2] Coren, S., Girgus, J.S., Erlichman, H. and Hakstian, A.R. An empirical taxonomy of visual illusions. Perception & Psychophysics 1976, vol 20(2), 129-137.

[3] Eco, U. On Beauty. Quercus Publishing MacLehose Press, 2010.

[4] Eco, U. On Ugliness. Rizolli, 2011.

[5] Internet World Stats. 3.035.749.340 Internet users estimated for June 30, 2014. www.internetworldstats.com/stats.htm, downloaded June 1, 2015.

[6] Kolb H, Lipetz LE. The anatomical basis for colour vision in the vertebrate retina. In: Gouras P, editor. Vision and visual dysfunction. Vol. 6, The perception of colour. London: Macmillan Press Ltd.; 1991. p. 128-145.

[7] Leonard, H. There Will Soon Be One Smartphone For Every Five People In The World. Business Insider Feb. 7, 2013. www.businessinsider.com/15-billion-smartphones-in-the-world-22013-2, downloaded June 1, 2015.
[8] Lepsius, K.R. Denkmäler aus Aegypten und Aethiopien. Editions De Belles-Lettres, Switzerland, 1897, Part III, volume 7, 217.
[9] lostcityoflondon.co.uk/tag/st-pauls-cathedral, blog, downloaded June 13, 2015.
[10] March, B. A Note on Perspective in Chinese Painting. The China Journal, Vol VII (2), 1927, pp. 69-72.
[11] Psychology Directory. psychology.about.com/od/jindex/g/just-difference-threshold.htm, downloaded June 13, 2015.
[12] publicdomainreview.org/collections/spring-morning-in-the-han-palace-17th-c, downloaded June 13, 2015. Creative Commons Attribution ShareAlike 3.0 license.
[13] Schacter, D.L, Gilbert, D.T., Wegner, and D.M. Psychology. Worth Publishers, New York, 2011.
[14] storiadellarteallempedocle.wikispaces.com. downloaded June 9, 2015. Creative Commons Attribution Share-Alike 3.0 License.
[15] Tuck, M. Gestalt Principles Applied in Design. sixrevisions.com/web_design/gestalt-principles-applied-in-design/, downloaded June 13, 2015.

3DIM: An Interactive 3D Map to Visualize Geo-Spatial Data in Mobile Devices

Ragaad AlTarawneh, Carl S. Marshall, Selvakumar Panneer,
Cindy K. Chung
Intel Labs, Hillsboro, OR, USA
{ragaad.altarawneh, carl.s.marshal,
selvakumar.panneer, cindy.k.chung}@intel.com

Abstract. We introduce an interactive 3D mobile-based map interface to visualize users' activities on Flickr. We designed the interface to handle limitations of smart devices' screen sizes by utilizing the new advancements in the performance domain of these devices, like relying more on GPU in rendering or producing high resolution pictures in real time. We compare the Flickr 2D map and the 3D interactive map of our app when they are rendered in smart devices. Our approach helps users keep the context of the picture and its geo-location without having to open a new view to explore the pictures and to get more details. We conducted a brief evaluation study where users preferred our proposed 3D interactive map compared to the 2D map, mainly due to the free rotation and the natural interaction techniques offered by our approach.

1 Introduction

Smart mobile devices (i.e., smartphones or tablets in our context) are now used for achieving many daily life routine tasks. Accordingly, their usage has increased dramatically over the last few years. However, ,many factors (such as limited screens sizes, new interaction styles, mobility, etc.) compared to desktop computing environments make it a big challenge to visualize big data on these smart mobile devices [1, 2, 3, 4]. Furthermore, handling and processing of various and possibly huge amounts of data (especially in the form of pictures) on these smart mobile devices is a growing challenge [2, 4].

In this regard, an interesting research question is how to adapt the visualization of big data into these mobile devices that on one hand support relatively limited capabilities (like screen size or battery life), while on the other hand, they support rich and natural interaction styles (e.g., multi-touch gestures). There have been many suggested optimization techniques to increase the battery life. However, we argue that the focus needs to shift towards using the Graphics Processing Unit (GPU) in the rendering phase rather than the standard Central Processing Unit (CPU). This would enable

developers to design their frameworks to utilize the parallel nature of GPUs, which would increase rendering speed and visualization performance dramatically.

With the evolution of current smart mobile devices more and more data are collected and stored, which open doors for new use cases and applications. For example, many social media tools enable users to share their data, pictures, or events. Facebook, Flickr and Instagram are few examples of heavily used platforms that offer these capabilities in real time. Due to the heavy usage of these platforms, they accumulate relatively big data over time. So, one challenge with such big data is how to summarize and present feature-rich data sets so that users/vendors can discover insights about their data more efficiently and effectively.

In this paper, we describe an interactive 3D mobile-based map interface that enables users of social media tools (Flickr[1] is a case study in our work) on smart devices to trace the most recent activities in the underlying social network. We describe the development of a picture-rich social media API viewer for the investigation of cultural trends of geo-tagged, user-generated pictures. We showcase the potential insights that could be gained from cross-cultural assessments in big data visualizations.

Finally, given the rich features in mixed media, we describe the advantages and limitations of the 3D interactive map visualization compared to the traditional 2D visualization on mobile devices. Our app allows real-time keyword queries of social media (i.e., Flickr in our case) and visualizes the pictures based on their geo-location tags. Users can interactively explore the 3D world to discover likeness or differences in trends of keywords across various regions. Comparing pictures of a keyword provides an immediate visual comparison of the term's popularity, as well as the relative conditions and behaviours associated with the term across regions.

The remainder of the paper is structured as follows: In Section 2, we describe our approach of visualizing the relations regarding the social media events and their geo locations. In Section 3, we discuss the result of a pilot study with 3 experts and we provide a comparison between the 2D map visualization and the 3D map visualization on mobile devices based on those experts feedback. In Section 4, we present the conducted user evaluation study and the overall results. Finally, we present our conclusions in Section 5.

---

[1] https://www.flickr.com/

2 Visualizing the Relations between Social Media Events and their Geo-locations

The very concept of designing a visual interface targeting big data in mobile platforms directly triggers the scalability issue [4], as it is required to scale up or down the visual interface according to the underlying device size. At the same time, it is quite a challenge to keep the intuitiveness of the design regarding the clarity, readability and interaction styles provided by such interfaces [3].

In this paper, we tackle the problem of designing an interface that encodes the relations between posts on social media (i.e., Flickr in our case) with their geo-locations. Although this problem has been tackled in Desktop platforms and for large screen sizes, relatively few studies have addressed the problem from the perspective of mobile devices' interfaces.

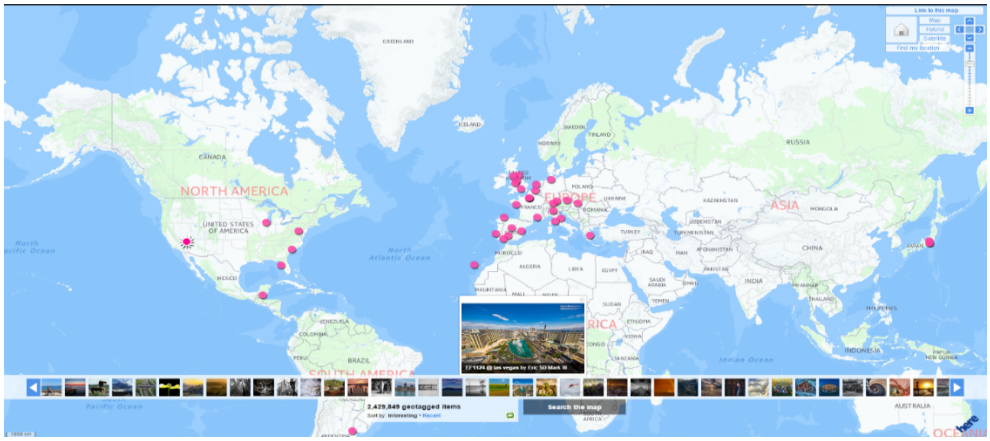


Figure 1: 2D Flickr map to show the pictures distribution over the world map, https://www.flickr.com/

The Flickr platform currently provides a 2D map interface as shown in Fig. 1. This interface works smoothly over desktop platforms, which are supported by a mouse interaction technique and relatively large-sized displays. Through Flickr's website, users can query publicly available Flickr's picture database with a search term, and obtain a 2D interactive map with dots to represent where relevant (defined by Flickr) pictures originated. The pictures are represented in a horizontal line at the bottom of this interactive map (see Fig. 1). Clicking on a particular dot on the map highlights the corresponding picture with caption at the bottom of the screen. Double-clicking takes the user to a full screen of the picture with picture owner information, caption, comments, likes, etc. The back button returns the user to the larger map. The pictures at the bottom horizontal bar of the screen can also be clicked, which results in highlighting of the corresponding dot on the map. Users can also

scroll the horizontal picture bar to discover a new set of pictures appearing in different geographical regions.

The main concern is the scalability and the readability of such a visualization on mobile devices that have relatively small screen sizes. To tackle this concern, we propose to replace the 2D world map with a 3D map representing the globe. In this case, the globe is rendered as a sphere textured with the world map; therefore, the cost of such visualization is not complex compared to the 2D map according to the GPU's rendering of it. This representation also offers a free rotation interaction style, where users can explore or rotate the globe using the swipe gesture interaction technique. We call this visualization the 3D Interactive Map (3DIM).

(A) (B) (C)

Figure 2: The 3D Interactive Map (3DIM) interface shows the relations between pictures corresponding to a key word search, and their geo-locations over the globe. (A) Shows the selected picture and aligns it with its location (in the United States), in addition to its caption. (B) Shows the other places that shared this picture's key-word using the color-coded lines. (C) Illustrates the idea of applying the 3DIM in mobile devices.

Our 3DIM approach provides an interface to visualize the relations between social media events and their geographical sources (see Fig. 2). The concept is very similar to the one provided by the Flickr 2D map visualization, where a user can query the database with search terms or click pictures to find their locations on the map. However, in our case, pictures from the search results are shown as discs on the geo-locations from which they originated, so that the point of the origin is immediately apparent with the most relevant visible pictures as defined by Flickr. The sizes of the discs represent the number of pictures from that particular region. This 3D effect allows for an immediate grasp of the popularity of a picture.

The 3D view of the pictures allows regions of the pictures to be visible while subtlety showing picture locations that would be occluded by the world. This presentation invites users to discover data trends by exploring world. By

clicking a picture disc, the user can see the caption of the picture. In addition, the comments of the picture become visible through color-coded lines that arch around the globe to the comment. Users can rotate through all the comments of a particular picture (see Fig. 2).

3 2D vs. 3D Map Visualization on Mobile Devices

We performed an initial pilot study with three experts with a visualization and interaction design background, in which we instructed them to explore both interfaces, i.e., the Flickr 2D map interface and our 3D map interface, on a windows based mobile device.

3.1 The 2D Interactive Map

As our experts tested the 2D map representation provided by Flickr on a smart device display, we observed the following:

The 2D world map was rendered over the whole display, which made it difficult to read as the map size was reduced due to the display size. The dots that represented the origins of the searched pictures were very small, which required additional interaction techniques to explore them properly. Also, the number of steps that were required to navigate through the picture database and to link them with their actual locations was relatively high; users had to identify the location through panning and swiping gestures followed by zooming in/out using the tapping/pinching gestures according to the location on the map. This process could lead to losing the context or the link between the selected picture and its location due to the change in the angle view.

3.2 The 3D Interactive Map

As our experts tested our interactive 3D map interface on a smart device, we observed the following:

The experts were able to interact with the interface naturally through the swipe gesture that was used to rotate the globe as well as to rotate the pictures surrounding the globe. Those pictures were distributed based on their original geo-locations. Tapping on a picture of interest increased the selected picture's size in order to show further details. In this case, the globe was rendered in the background. This enabled the experts to link between pictures and their sources without losing the context. The interface also showed the set of links between a selected picture and the locations where users in those locations also shared the same picture. In this regard, we didn't notice any occlusion

resulting from the 3D effect, as the globe was rendered in transparent fashion when a picture was selected. We did not test the rotation facility using the accelerometer sensor facility.

3.3 Discussions

Upon comparing the Flickr's 2D visualization and our 3D map visualization, we noticed some immediately apparent differences, e.g., the representation of query results across geographical locations or the differences in construal for each query. Some examples of queries that were run to compare the 2D and 3D visualizations include:

- *Cat:* A popular focus on the social media, the "cat" query produced a large number of pictures across the world. It was observed that there are relatively larger cats (e.g., lions, cheetahs, etc.) in the open wild in Africa while relatively more domestic cats in other areas of the world. This fact was instantly more apparent on the 3D map compared to the 2D map.
- *Nature:* Vegetation and landscapes naturally vary by geographical conditions and this was apparent in both the 2D map and the 3D map. However, the experience of the climate and vegetation continuity and differences across regions were more immediately apparent in the 3D version.

From our experience, we observed that using the 3D map on mobile devices saves more space compared to the 2D map that expands horizontally and vertically over the device display. Users have more freedom and more intuitive interaction styles due to the free rotation interaction of the 3D map. However, the 3D map visualization was designed to show extra data only on demand, unlike the 2D map. For example, in the 3D map, the selected picture was shown and enlarged only compared to the other pictures upon user selection. Further, once the selected picture was enlarged it sometimes occluded the background, which hindered the readability or ability to detect relations between the selected picture and the other places that shared this particular picture.

4 The User Evaluation Study

We conducted an explorative user evaluation study, in which we invited 12 users explore and comment on both map visualizations, i.e., the 2D map visualization by Flickr and our 3D map visualization. All participants had a computer science background. Half of the participants were females, and their

ages ranged from 23 to 56 years. Users were selected randomly and were encouraged to participate in the study freely and voluntarily.

At the start of the experiment, an introduction to both platforms was given to each participant [by voice/experimenter/online?]. Half of the participants were randomly assigned to start with the Flickr 2D map visualization and then our 3D map visualization, while the other half did it in the reverse order. Both visualizations were rendered on a Windows-based tablet. Users were briefly introduced to how to use both apps and shown what their interesting features are. Then, they were given 5 to 10 minutes to play with both maps and to explore the features by themselves. After this, we asked them to give us their feedback through a closed-ended questionnaire form.

Figure 3 shows users' feedback and their preferences that we gathered from the given set of questionnaires. From the results, we can see that both map representations have the same number of votes, i.e., 6 votes for each, when users were asked which map they would like to use in the future. Results indicate that users found the 3D map visualization more intuitive, as it requires fewer interaction steps compared to the 2D map. Moreover, users found that the 2D map visualization was more readable and used the space more efficiently. However, this might be because the test was done on a tablet device. This feedback could be different in the case of using a smartphone with smaller screen.

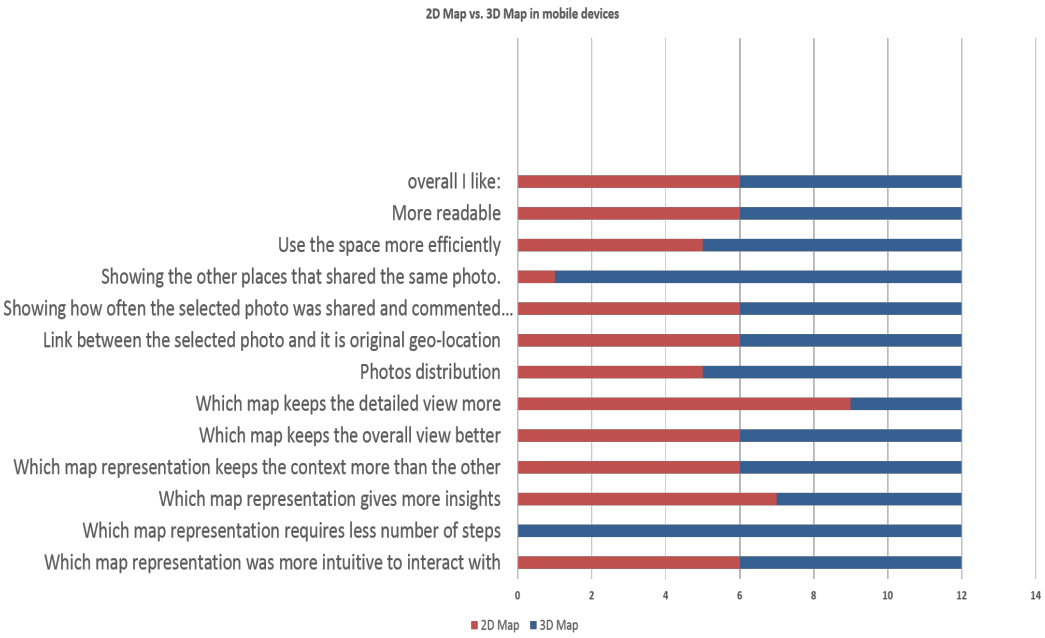


Figure 3: Overall results showing users feedback upon using 2D map and 3D map.

5 Conclusion

We presented our approach of visualizing 3D geo-spatial data on smart mobile devices. More specifically, we visualized Flickr pictures across their geo-location origins. For this, we proposed using a 3D map instead of the 2D map that has already been developed by Flickr. Our interactive 3D map utilizes the space on limited mobile screens more efficiently compared to the 2D map and enhances the grasp of the pictures' geo-distribution faster and clearer.

We compared the Flickr 2D map visualization and our 3D interactive map using mobile devices and compared the interaction steps that users needed to query some aspects on the visualizations. In summary, differences in cultural construal of concepts were more immediately apparent and more information was conveyed through the 3D effects without overwhelming the user. The 2D visualization required more interaction steps to dig deeper into the data although the general locations in which a search term was located or where it simply did not exist were more immediately apparent.

In future work, we will consider the users' feedback in optimizing our 3DIM map App. Moreover, we will investigate how to allow users to interact with this App utilizing multi-touch schemes provided by tablet devices.

References

[1] G. Abowd, C. Atkeson, J. Hong, S. Long, R. Kooper and M. Pinkerton. Cyber guide: a Mobile Context Aware Tour Guide. Wireless Networks, 3(5), 421- 433, 2000.

[2] Tiago Gonçalves, Ana Paula Afonso, Maria Beatriz Carmo, and Paulo Pombinho. Overview "vs" detail on mobile devices: a struggle for screen space. BCS-HCI '12, British Computer Society, Swinton, UK, 233-238, 2012.

[3] S. Burigat and L. Chitarro. Visualizing the Results of Interactive Queries for Geographic Data on Mobile Devices. GIS '05, ACM, 277-284, 2005.

[4] Paulo P. Pombinho. 2010. Information visualization on mobile environments. MobileHCI '10, ACM, New York, NY, USA, 493-494, 2010.

Simplifying the Input of Perceived Exertion in the Mobile Context using Prediction

Janko Timmermann[1], Alexander Schiotka[2], Wilko Heuten[1], Susanne Boll[2]
[1]OFFIS – Institute for Information Technology, Oldenburg, Germany
[2]University of Oldenburg, Germany
firstname.lastname@{offis.de, uni-oldenburg.de}

Abstract. In cardiac rehabilitation, training in the right intensity is important. To find this right intensity, vital parameters can be measured, but also the perceived exertion is an important factor. The perceived exertion is often measured using the Borg-RPE-Scale. It is usually used by trainers in indoor environments, not in a mobile context. In this context, the complexity of the Borg-RPE-Scale is a problem, especially when it should be used on a mobile device. In this paper, we present an approach to predict the perceived exertion using the vital parameters of a user to enable the realization of simple user interfaces. We show that the prediction and therefore the simplification is possible. Our results indicate that the usage of a predicted value seems not to influence users when entering their perceived exertion.

1 Introduction

Cardiovascular diseases are an enormous threat for our modern society, responsible for nearly 25 percent of all deaths in the world [10]. For patients who already suffer from a cardiovascular disease and probably already had a heart attack, the cardiac rehabilitation is a very important mean to achieve the best possible recovery [2,3,4,5].

The core element of cardiac rehabilitation is physical activity, which helps to improve the overall constitution of the patient. However, the right intensity is hard to determine because of many important influencing factors like age, constitution, weight, health, and medication [1].

Measuring the heart rate is the most common way to determine the training intensity. However, such measurements might not be available sometimes or be inaccurate, e.g. due to arrhythmia. Another way to assess the training intensity is to ask the patients about their perceived exertion. It is an important parameter as it reflects the subjective health of a patient. It helps to understand the patient's sense for exertion and further allows adjusting the training intensity in a way the patient is comfortable with. Quantitative scales have been developed to overcome the subjective nature of exertion. In cardiac

rehabilitation, the Borg-RPE-Scale is used most frequently. In practice, trainers verbally ask patients about their perceived exertion and enter the value on the actual scale. The training is then adopted appropriately. While this is fine for supervised indoor training, patients in unsupervised outdoor training are asked to record the value on their own without any assistance, which can be quite hard. To stop and enter the required information is not an option, as this potentially changes the perceived exertion and patients might lose their preferred training pace. On mobile phones, which are more and more used to accompany the exercise, existing interaction paradigms, like a drop down menu, often require several corrections until the right exertion value is recorded. Consequently, perceived exertion is hardly recorded in outdoor training, which limits the outdoor trainings' success and puts the patients' well-being at risk.

In this paper, we investigate a system which uses an algorithm to predict the Borg-RPE-Points of the patient using measured heart rate and breathing frequency. This way, the system can generate an estimated value to simplify the entry of the real value for the user. We evaluated the system in a user study, to test for possible influences due to a predicted value. We conclude that our system is able to predict the value with an appropriate accuracy and that influences are unlikely. We also show indications on how a future system should look like and how the mobile usage of the Borg-RPE-Scale can be realized.

2 Background

For patients with a cardiovascular disease it is important to perform physical activities with the right intensity, duration, and frequency. The required intensity strongly varies depending on e.g. age, weight, illness, and medication [1]. One indicator for a good intensity is the perceived exertion. To measure the perceived exertion, the Borg-RPE-Scale (RPE: "Ratings of Perceived Exertion") was developed by Gunnar Borg [1]. The Borg-RPE-Scale is used nowadays to measure the perceived exertion of patients [6]. The Borg-RPE-Scale ranges from 6 to 20, so 6 means no exertion at all and 20 is the maximum exertion a patient can experience. The Scale is depicted in Figure 1. An activity at that level can only be done over a short period of time and is experienced as the highest possible load level.

Numerical	Description	Graphical
6		
7	Easy	
8		
9	Very Light	
10		
11	Fairly Light	
12		
13	Somewhat Hard	
14		
15		
16	Hard	
17		
18	Very Hard	
19		
20	Maximum	

Figure 1. Borg-RPE-Scale

Though the perceived exertion as response to physical activity shows an exponential increase [1], the Borg-RPE Scale is linear. Using the Borg-RPE-Scale, patients can report their perceived exertion during their training. This information can then be used for reflecting the training, adjustments regarding the intensity, or even as main parameter for training adaption [11], so the patient can do her or his training in a comfortable way. Technically, the reports can be used to verify the heart rate measurements and allow a support system to adapt to users where the heart rate cannot be measured appropriately or does not normally reflect the intensity due to arrhythmia or medication.

Many parameters can influence the perception of the exertion. Especially patients with cardiovascular diseases sometimes have a deranged perception, but also environmental factors like e.g. temperature, humidity, and noise can influence the rating. Medication can also lead to a deranged perception and to different load limits. Under optimal circumstances, the points on the Borg-RPE-Scale and heart rate show a strong correlation between 0.8 and 0.9 [7]. This correlation is unrelated to the type of the physical activity [3]. For healthy people who are between 30 and 40 years old, the Borg-RPE value can be computed by dividing their heart rate by ten [1]. Patients who are treated with beta blockers have the same perception of exertion as without the medication, though their heart rate is lower [7].

3 Context Analysis

To understand the usage scenarios for the Borg-RPE-Scale during cardiac rehabilitation, we conducted a semi-structured interview with two trainers from a heart clinic and rehabilitation centre. The leading questions for the

interview were: "*Which vital parameters correlate with the perceived exertion?*", "*Is the Borg-RPE-Scale already used in outdoor training?*", and "*What is the time interval for measuring a patient's perceived exertion?*" The trainers reported that the Borg-RPE-Scale is successfully used during cardiac rehabilitation to assess exertion and to optimize the training intensity for the patient. The trainers agreed to the fact that medication can strongly influence the perceived exertion and are therefore a problematic factor for its assessment. Another problem we did not see in related work is possible shame of the patient. Some patients do not like to tell their perceived exertion and report a wrong value which seems to be more convenient for them. Especially men sometimes tend to report smaller Borg-RPE values so it seems they have more capacities than they actually have. Since there was no experience with assessing the perceived exertion not through a trainer but through a technical device, it was unclear if that effect is also relevant in a mobile scenario. The trainers were optimistic about the usage of the Borg-RPE-Scale during outdoor activities. The time span between two measurements should be more than 6 minutes to let the cardiovascular system adjust itself to the changed exertion. An interval of 15 minutes was considered practically useful. Additionally, an assessment should be done if a known change of environmental factors, like changing slope, happens because it is very likely that a change in perceived exertion would also happen then.

In preliminary tests, we tested the input of the perceived exertion on the Borg-RPE-Scale with the help of common input methods on mobile devices, i.e. drop down boxes, sliders, and buttons. All variants needed a big and high resolution display to be easily used. Voice input would be an option for entering the perceived exertion and may be the best choice for some users. However, for many users the use of voice input in public is still awkward and the problem of misunderstandings is quite relevant especially in public and loud environments.

The input speed is critically important, since the input distracts users from their activity. Thus, a simple method for entering the value is needed. We propose the usage of a predicted value which requires only small adjustments by the user to ease the user input.

4 Approach

As our approach to simplify the usage of the Borg-RPE-Scale in the mobile context during the cardiac rehabilitation, we created an algorithm which predicts the users perceived exertion as Borg-RPE-Points using the measured heart rate and breathing frequency as the two main input parameters. As

explained in the background section, the heart rate shows a very high correlation to the Borg-RPE-Points in an optimal scenario. We used the breathing frequency as a supportive parameter since it can be easily measured alongside the heart rate and is also related to the exertion. As third parameter, the correction to previously computed values the user entered is taken into account to adapt the computation to the specific user.

The outcome of using the algorithm is the possibility to use input methods where only minor corrections have to be entered by the patients. Thus, small wearable devices which can be worn e.g. at the wrist can be used instead of regular and big smartphones, which have to be hold in the hands or need to be pulled out of a pocket to enter the perceived exertion. The algorithm is not intended to convert the Borg-RPE values into a measured value. It is important that such an algorithm helps users to report their perceived exertion, but in no way performs the rating for them.

The algorithm computes the estimated Borg-RPE value using the heart rate and breathing frequency as vital parameters. These parameters have to be measured by sensors the user is wearing. Further, the algorithm needs information about the minimum and maximum heart rate. Because sometimes users do not know these values, we used the popular formula 220 – age. Very high accuracy was not an important factor for the algorithm, because it was not supposed to give recommendations about the load level to the user. The resting breathing frequency is always assumed to be 12 breaths per minute if it's not entered by the user [8]. For trained users, the algorithm assumes at maximum 60 breaths per minute, otherwise 45 breaths per minute [9] are assumed.

The algorithm starts with computing the Borg-RPE value for each vital parameter. As described in the background section, the Borg-RPE value is derived from the heart rate and can be approximately computed by dividing the heart rate by ten. Instead of doing only this simple operation, we took the minimum and maximum heart rate into account and divided this range into 15 sections which represent the 15 values of the Borg-RPE-Scale. If the actual heart rate exceeds or deceeds these limits, the entered or computed maximum or minimum values are used instead. The same method was used to map the breathing frequency onto the Borg-RPE-Scale. To combine them, we calculate the average value of both. Thus, both values have the same weight in the final computed value.

To adjust the computation to a specific user, the algorithm also takes the difference between this computed value and the corrected value of the user into account. The algorithm computes the average of all these corrections and applies this offset to the computed value.

We created an Android application to test our approach and collect user feedback about the Borg-RPE-Scale on mobile devices. The application was kept simple and was planned as throw-away prototype. It was not tested for usability or designed using dedicated guidelines since user interaction with the application was supposed to happen only supervised. It consisted of two views: The start view and the training view. The start view let users enter their gender, training level, and age (Figure 2). An alternate tab let users enter their resting and maximum heart rate and breathing frequency. Both had a button to control whether the application should start the training with computed Borg-RPE value suggestions or not.

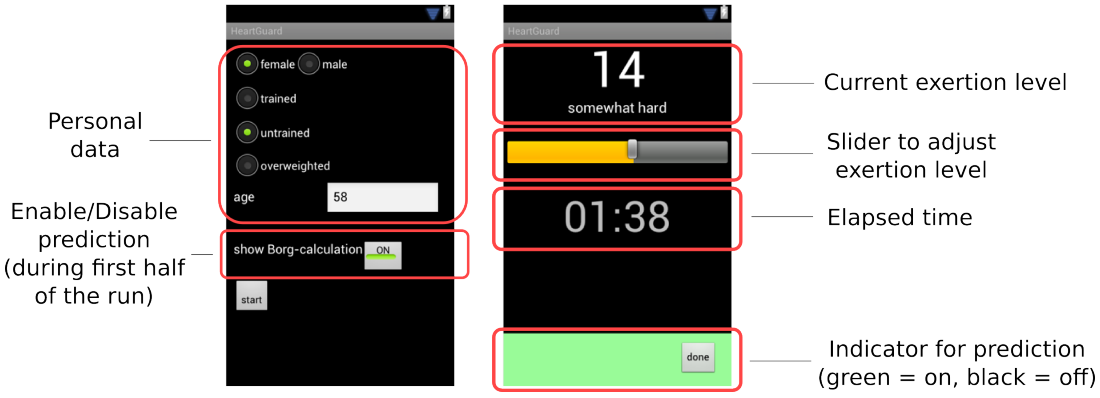


Figure 2. Start & Training View of our application.

The training view was shown after the start view. It is depicted in Figure 2. On top of the view the selected exertion level on the Borg-RPE-Scale as numeric value was shown. A description of that level was printed below in text form. To change this value, a slider was placed below this display. The slider let users enter all possible values from six to 20 by moving it from left to right. It was not only controllable by the users, but also by the algorithm to keep the exertion level display and the slider always consistent. A timer showed the time since the training was started. Near the finish button to end a training, the bottom area of the view was coloured black or green depending on whether the algorithm to suggest values was active or not. Intentionally, the app did not show the measured heart rate and breathing frequency to avoid influencing the users' perceived exertion input.

5 Study

To test the algorithm for predicting the Borg-RPE-Points, to evaluate whether a displayed predicted value influences the users, and to collect qualitative

feedback about a mobile usage of the Borg-RPE-Scale, we conducted a field study.

The location for the study were paths dedicated for running. Thus, no distraction by traffic was possible and the participant was able to run continuously. In most cases, the location was a sports field with a running track. We used a Samsung Galaxy Nexus running the application described above in combination with a Zephyr BioHarness 3.0 belt. The BioHarness 3.0 sensor allows measuring the heart rate and the breathing frequency.

We conducted the field study with 14 participants, 4 female and 10 male. The age ranged from 22 to 44 years. The mean age was 26.1 (SD: 5.6). All participants were healthy, despite the fact that the target user group are patients with a heart disease. We did this due to ethical reasons, especially to avoid any risk for the participants. However, the algorithm regards special requirements of cardiac patients like adapting to a lowered heart rate due to medication. As described in the context analysis section, these users have the same perceived exertion compared to healthy users and the algorithm adapts to their lower heart rate over time. The activity levels of the participants were quite mixed and reached from no physical activity to regular intense physical activity. The participants were not paid for their participation in the study.

The study consisted of three parts with different load intensities. The intensities were *low (walk fast), medium (run with moderate pace), and high (run fast)*. The order in which a participant had to reach these load levels was defined using the Latin Square. Every intensity was meant to be maintained for ten minutes. If a participant was not able to maintain a load level for such a long time, she or he should try to keep the highest possible load level. As mentioned above, the application was able to query the load level from the user with or without supposing a predicted value. This behaviour switched in the middle of a run, thus after five minutes. For the first run of the first participant, it was randomly defined that no predicted values were shown in the first five minutes but thereafter. For all following runs of all participants, this order was switched for each run. After one run, a short break for the participants was planned to normalize the heart rate and breathing frequency before starting the next run.

Afterwards we conducted a semi-structured interview with the following leading questions: "*What do you think about the application?"*, "*Were you satisfied with the supposed Borg-RPE value and was there a difference between different load levels?", "Did you like this type of representation of the exertion?*", and "*Would you like to use such a system?*"

In the beginning of the study, we met each participant at a location communicated beforehand. We explained the study to them and answered

their questions. After the participants signed the informed consent, we explained the equipment which we described above. We handed the equipment to the participants and let them attach the chest strap in a private environment. Afterwards, we explained the Borg-RPE-Scale and ensured the participants understood the principle of reporting the perceived exertion with it. Finally, we started the three runs with altering intensity and with pauses in between. The length of a pause was defined by the participants but was encouraged to be only as long as needed to recover from the previous run and normalize the heart rate and breathing frequency. In the end, we conducted the interview, thanked the participant, and ended the study.

6 Results

The quantitative results are based on the data of twelve of the 14 participants. Two recorded data sets contained very high and obviously erroneous recorded heart rates and were therefore eliminated for this evaluation. The following table shows the means of our recorded values during the study at the different load intensities:

	Low Intensity	Medium Intensity	High Intensity
Heart Rate	144 (SD: 17.2)	180 (SD: 10.9)	188 (SD: 11.3)
Breathing Frequency	31 (SD: 5.7)	35 (SD: 6.8)	39 (SD: 4.7)
RPE Ratings	14 (SD: 2.0)	17 (SD: 1.8)	18 (SD: 1.3)

Table 1. Means of values recorded during the study

The correlation coefficient (Bravais-Pearson) between the heart rate and the Borg-RPE-Points was 0.8, as in the related work. The breathing frequency showed a medium correlation of 0.47, which is still good, but clearly shows that the breathing frequency should only be used if the heart rate is not available or inaccurate.

The average ratings of perceived exertion showed no significant difference (t-test, $p = 0.69$), when the suggestions by our algorithm were enabled (13.97, SD: 2.84) or disabled (13.85, SD: 3.47). However, the count of adjustments the users made to the predicted or previously entered value differed significantly (t-test, $p < 0.01$): The displayed Borg-RPE value was corrected on average 8.1 times (SD: 2.68) of 15 when the algorithm was used and 5.28 times (SD: 2.67) of 15 when not. If users corrected the value, it was changed

on average by 2.11 Borg-RPE-Points (SD: 1.07) when the algorithm was used and 1.2 Borg-RPE-Points (SD: 1.16) when not. The Shapiro Wilk test showed that this data set was not normally distributed ($p < 0.01$), so we used a Wilcoxon signed rank test as test for significance. The test showed that the difference is not significant with $V = 73$ and $p = 0.06$. These results show, that the quality of the algorithm is not well enough in this iteration. However, it is also an indicator that the algorithm did not influence the users, since they did not simply use the predicted value instead of correcting it. The fact that the average exertion did not significantly differ between both conditions also supports this indication. Under the assumption of no influence we can recommend the usage of the Borg-RPE-Scale with the help of a predicting algorithm in the mobile context, if parameters for assuming values in an accurate way are available. A future version of a prediction algorithm should only propose a predicted value, if the exertion changed heavily, otherwise it may make the input of the Borg-RPE value even more complicated than easier. Naturally, if the predicted value differs not so much from the previously entered value of the user, the proposal of the predicted value instead of the previously entered value would not be beneficial.

During the interview at the end of the study, we found that users would like to use a system to record their perceived exertion if the interface is easy to handle during physical activity. Most users were satisfied with the suggested value given by the app, but reported a higher accuracy at higher load levels, which is supported by the quantitative results.

7 Conclusion

In this paper, we tested the usage of the Borg-RPE-Scale on mobile devices. We created an Android application which queried the Borg-RPE value from the user during training. This application was combined with an algorithm to predict this value using heart rate and breathing frequency to simplify the user input of the perceived exertion. We found that our algorithm was able to predict the perceived exertion within an acceptable range and is therefore able to enable a user input with only small corrections of a predicted value. Our results indicate that users were not influenced by these predictions. We can therefore recommend the use of this technique if the users are intended to record their perceived exertion using the Borg-RPE-Scale. However, it has to be made clear to the users, that a predicted value does not replace entering the perceived exertion but eases the input. We found that users were generally interested in a system which enables the input of perceived exertion. The predicted exertion level of the application was well accepted. When using the

Borg-RPE-Scale in the mobile context, the prediction algorithm needs to be optimized to raise the accuracy. It should only be active when the activity clearly changed and the Borg-RPE value is expected to raise or fall much. Further input methods have to be created and evaluated which are possible with the help of the algorithm. This includes also the usage of small displays like smartwatches or other wearable devices. Since users were interested in using the history of their perceived exertions as observation for their training, this has to be tested against traditional systems, like heart rate recorders.

References

[1] Borg, G. Borg's perceived exertion and pain scales. Human Kinetics Publishers, 1998.

[2] Franco, O., De Laet, C., Peeters, A., Jonker, J., Mackenbach, J., and Nusselder, W. Effects of physical activity on life expectancy with cardiovascular disease. Archives of internal medicine, vol. 165, no. 20, p. 2355, 2005.

[3] Gamberale, F. Perceived exertion, heart rate, oxygen uptake and blood lactate in different work operations. Ergonomics, vol. 15, no. 5, pp. 545–554, 1972.

[4] Ignarro, L., Balestrieri, M., and Napoli, C. Nutrition, physical activity, and cardiovascular disease: an update. Cardiovascular research, vol. 73, no. 2, pp. 326–340, 2007.

[5] Leon, A., Franklin, B., Costa, F., Balady, G., Berra, K., Stewart, Thompson, P., Williams, M., and Lauer, M. Cardiac rehabilitation and secondary prevention of coronary heart disease an american heart association scientific statement from the council on clinical cardiology (subcommittee on exercise, cardiac rehabilitation, and prevention) and the council on nutrition, physical activity, and metabolism (subcommittee on physical activity), in collaboration with the american association of cardiovascular and pulmonary rehabilitation," Circulation, vol. 111, no. 3, pp. 369–376, 2005

[6] Löllgen, H., and Ulmer, H. Das gesprärch während der ergometrie: Die Borg-skala." Deutsches Arzteblatt, vol. 101, pp. A1014–1015, 2004.

[7] O'Sullivan, S. Perceived exertion a review. Physical Therapy, vol. 64, no. 3, pp. 343–346, 1984.

[8] Schein, M., Gavish, B., Herz, M., Rosner-Kahana, D., Naveh, P., Knishkowy, B., Zlotnikov, E., Ben-Zvi, N., Melmed, R., et al. Treating hypertension with a device that slows and regularises breathing: a randomised, double-blind controlled study. Journal of human hypertension, vol. 15, no. 4, pp. 271–278, 2001.

[9] Stevens, E., and Randall, D., “Changes in blood pressure, heart rate and breathing rate during moderate swimming activity in rainbow trout. Journal of Experimental Biology, vol. 46, no. 2, pp. 307–315, 1967.

[10] Taylor, R., Brown, A., Ebrahim, S., Jolliffe, J., Noorani, N., Rees, K., Skidmore, B., Stone, J., Thompson, D., and Oldridge, N. Exercise-based rehabilitation for patients with coronary heart disease: Systematic review and meta-analysis of randomized controlled trials. The American journal of medicine, vol. 116, no. 10, pp. 682–692, 2004.

[11] Timmermann, J., Workowski, A., Heuten, W., Willemsen, D., and Boll, S. Human centered training: perceived exertion as main parameter for training adaption. NordiCHI 2014. ACM, 2014, pp. 995–998.

[12] W. H. Organization. (2011, June) The top 10 causes of death. World Health Organization. Last visited: 11th December 2012. [Online]. Available: http://www.who.int/mediacentre/factsheets/fs310/en/index.html

Smart Ecosystems and the Impact on Mobile Interaction Design Methods

Claudia Nass, Konstantin Holl, and Christian Müller
Fraunhofer Institute for Experimental Software Engineering IESE
Fraunhofer-Platz 1, 67663 Kaiserslautern, Germany
{claudia.nass,konstantin.holl,christian.mueller}@
iese.fraunhofer.de

Abstract. Smart ecosystems combine the Internet of Things, Services and Data by connecting several system classes. A crucial player are mobile apps aimed at covering the mobility requirements of the individual participants of such smart ecosystems. The characteristics of mobile apps call for tailored design approaches (e.g., addressing the complexity and size of such systems or the seamless integration in the daily life of people). With regard to the creation of the interaction design, we have identified challenges arising in the context of smart ecosystems and analysed some representative mobile interaction design approaches that can match these challenges.

1 Introduction and Motivation

The penetration of mobile devices in the daily life of people is growing rapidly [1]. People are using their mobile devices for shopping, communicating, and for accessing information more than ever. In some regions of the world, mobile devices are the main media to access internet services [1]. Especially in the context of smart ecosystem, mobile devices plays a very important role. Originated in the biology, ecosystems describe a system resulting from interacting organisms and their environment where organisms try to achieve certain goals considering their environment. Hence, smart ecosystems are the technical equivalent connecting information systems and technical systems with their environment. An example of a smart ecosystem is the modern integration of "mobility solutions that are able to jointly use different means of transportation smartly in order to solve a given transportation problem" [2].

Since mobile apps are an essential part of smart ecosystems, the impact on the mobile interaction design needs to be investigated. In this paper, we present some main trends of smart ecosystems and their influence in the mobile interaction design. We also explore two questions regarding the impact of smart ecosystem in mobile interaction design methods: (1) the interplay between challenges of smart ecosystems and mobile interaction design as well as (2) the requirements of these challenges on mobile interaction design methods and approaches.

2 Trends of Smart Ecosystems and the Role of Mobile Devices

Smart ecosystems enable a new way of life, allowing people to have constantly access to a wide variety of new services and products. Different scenarios, as represented in Figure 1, are examples of smart ecosystems: an autonomous car that meets you in the street and drives you to an important meeting, where business partners are waiting for you. A smart ecosystem would be able, e.g., to recognize the appointment in your calendar, to offer you the best transportation option, and to autonomous drive you to the current position of your partners. This means a system that combines emergent software (network of information systems) with cyber-physical systems (network of embedded systems) [3].

This is enabled just by the strong **digitalization of society**. The increasing availability of the Internet allows individuals to be in touch with the rest of the world almost permanently and ubiquitously. People are integrated into ecosystem processes, delivering information and consuming services through their daily routine, especially by using mobile devices.

The knowledge gained from people due to their interaction with smart ecosystems integrated into enterprise processes enables **mass customization** of products and services. The use of customised advertising when people are physically close to a store, also known as *geofencing*, demonstrates how user behaviour (elicited from their use of consuming online services and from individual movement patterns) is used by enterprises within their business processes. The use of integrated information technology allows enterprises to "provide individually designed products and services to every customer through high process agility, flexibility, and integration", i.e., mass customization [2].

Process attributes such as agility, flexibility, and integration can only be achieved with **high levels of automation or even autonomous systems**. This means, e.g., that production and logistic systems are prepared to respond to new demands. An example is the use of drones for delivering products (e.g., medicine) quickly and autonomously to a person, meaning the customer himself is becoming the new delivery address provided by the sensors of the mobile devices.

Therefore, information systems, embedded systems, and individuals form a **highly networked system** in the context of smart ecosystems that supports the continuous interchange of information and smart functionalities. Mobile devices empower this complex system because they allow permanent integration of people into the smart ecosystem.

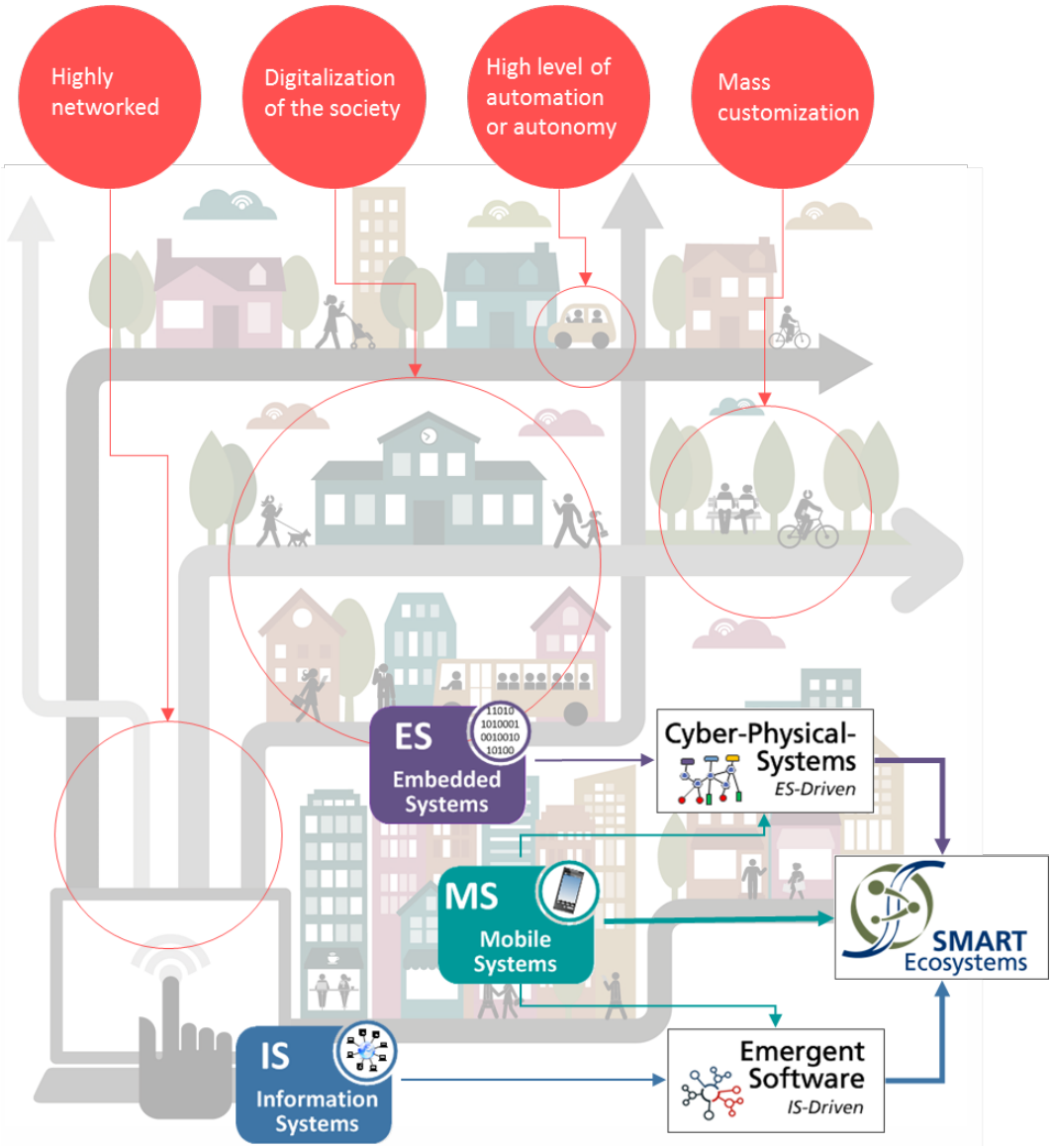


Figure 1. Smart Ecosystems and their Trends (adaption from Knodel and Webel [3])

3 Challenges of Smart Ecosystems and their Impact on Mobile Interaction Design Methods

Mapping challenges of smart ecosystems to common challenges of mobile apps beyond smart ecosystems requires first the consideration of general characteristics of mobile applications.

3.1 Mobile-specific Characteristics

Differences in developing and using mobile apps compared to classical software are often related to technological aspects. The main characteristics of applications on mobile devices such as smartphones and tablets according to Muccini et al. [1] are listed below, together with their impact on interaction design:

- **Limited resources**

The availability of energy and memory space, in particular, is lower for mobile devices than for classical desktop devices. This needs to be considered by the interaction design. For example, a permanently running feedback process has energy requirements that can be fulfilled by a desktop computer without any problems, but could lead to a battery drain of a mobile device that is unacceptable for the user.

- **User interface**

The main interface for the user interaction on current mobile devices is a gesture-based screen. It represents the primary input and output interface and offers interactions such as a variety of gestures using one or multiple fingers. However, the use of voice interaction and other sensors integrated in mobile devices cannot be neglected and has to be improved in high interactive scenarios. Methods for interaction design need to be aware of the range of interaction possibilities.

- **Context Awareness**

Besides the user interface, mobile devices use several sensors to capture context information depending on the environment of the device, such as brightness, connectivity, motion, or location. This context awareness can be involved in the user interaction and, as a consequence, it has to be considered as part of the interaction design.

- **Diversity**

The diversity of devices and operating systems results in high complexity of variants. Hence, a mobile app is often required to be compatible with different platforms using different interfaces. For the interaction design, this means that it has to handle aspects such as varying sizes of screens or even the availability of hardware buttons.

3.2 Challenges of Smart Ecosystems

Considering the mobile-specific characteristics together with the trends in smart ecosystems, presented in section 2, lead to a set of challenges that impact mobile interaction design. To answer our research question, we need to analyse if the current methods for designing mobile interaction are able to support the development of mobile apps in the concrete context of smart ecosystems, if there is a demand for adapting general methods for designing mobile interaction, or even developing new methods for the context of smart ecosystems.

The first main challenge presented by smart ecosystems involves the **complexity and size** of such systems. Nowadays, the design of mobile apps includes not just the interaction with the mobile app itself but also the whole experience with the service. Information volume, interaction intensity and availability, and cognitive load are aspects that have to be considered when designing new mobile interaction.

Additionally, we have the seamless integration of the real world with the digital world, i.e., the **integration of heterogeneous systems**. On the one side, there are information systems that process and control the interaction of the user and, on the other side, there are physical objects in which sensors and embedded components influence the interaction of people with their device (even mobile devices possess this characteristic).

Furthermore, the use of new mobile devices (e.g. smart glasses or watches, or shirts) forces us to think about their integration into people's daily lives, especially with regard to quality aspects as safety, privacy, intimacy, and fun. **Guaranteeing quality** attributes in this context is even more relevant due to the intensity of use.

Another challenge is related to **standards and regulations** for the communication between devices and embedded components. The development of communication protocols, interfaces and methods for easy integration is a challenge for mobile interaction design and development. Another important factor are ethical aspects that involve the intensive use of systems in daily life, such as user manipulation.

3.3 Impact on Mobile Interaction Design Methods

The challenges arising from smart ecosystems that are not just mobile-specific lead to further impacts on mobile interaction design.

Challenges in Smart Ecosystems	**Impact on Mobile Interaction Design Methods**
Complexity and size	• The complexity and the size of smart ecosystems cannot be detached. Methods should support designers and experts to identify and organise the volume of information that emerges with smart ecosystem (dominating the chaos). It has to consider the whole experience and service in which the user is participating. • Example: Breaking vast amount of information down into single tasks like when monitoring drone delivery. The user should be able to focus on essential functionality to take over the control.
Integration of heterogeneous systems	• User will still interact with real objects, not everything will work with information systems on smartphones or apps. Methods should support the development of new concepts of interaction, considering the focus of integrated sensors in daily objects as well as in the mobile devices itself. • Example: The integration of sensor for perceive the user situation

	(e.g. shirts with heart beat sensor) as well as displays integrated in common objects (e.g. smart cups) will enrich our interaction with smart ecosystem
Guarantee of quality	• Methods should support a systematic construction of quality aspects in the interaction design of mobile applications, both in the identification of context-based quality attributes as well as in the active integration of such qualities in mobile app design. • Example: The interaction design related to a drone that is part of a smart ecosystems leads to required quality guarantees e.g. even regarding safety and security.
Standards and regulations	• Methods should give clues for the integration with other smart ecosystems classes. Especially in the (short-term) development of mobile apps it is important to consider technical aspects, such integration in other smart ecosystems, in early phases of the design process. • Example: The interplay of mobile devices and devices likes drones leads to a clash of standards and regulations like regulated flight times that need to be considered by the mobile software.

Table 3: Challenges in smart ecosystems and their impacts on mobile interaction design methods

4 Coverage of Challenges by Existing Methods

To answer our research question, we need to relate the identified challenges and corresponding impacts (see Table 3) to the abilities of existing mobile interaction design methods. Table 4 presents some representative methods already used in projects by the authors and indicates the challenges faced by these methods when developing mobile apps in the context of smart ecosystems.

	Complexity and size	Integration of heterogeneous systems	Guarantee of quality	Standards and regulations
mConcAppt [6]		x	x	
mPotential [7]		x		
In Situ Design [8]		x		
A Diary Study of Mobile Information Needs [9]	x	x		
Experience Clip [10]		x	x	

Table 4: Coverage of challenges by existing methods

mConcAppt (*Conception of Mobile Apps*) is a framework that supports UX experts and developers in designing new mobile applications. The main value of mConcAppt is the systematic consideration of quality aspects in this user-centered method [6]. The framework focuses in the quickly generation of interaction concepts that support concrete task in business process. Because

the mobile apps are created in the context of a concrete process, we consider that this method attend partially the integration with other systems. An improvement potential in the framework is related to the consideration of more aspects of the user context when conceiving the mobile app. Particularly the use of a technique called *product philosophy* that considers quality attributes through the conception of mobile interaction design enables the integration of adequate quality attributes in the mobile app.

mPotential supports UX and requirement experts in identifying the mobility potential of processes and roles [7]. Using mPotential, it is possible to identify in the context of an organization the roles and tasks that have the potential to be supported with a mobile app. The method allows the detailed modelling of role, task, transit time, work time outside of the main work place and special physical conditions during the execution of tasks. This leads to the derivation of improvement potential ideas in terms of mobile apps.

In Situ Design provides techniques that enable requirements elicitation and user testing in the place of use of mobile devices [8]. This technique leads to a collaborative design of mobile apps in the place that it is used with the help of representative of the target-group. The main insights increase the integration of the mobile app in the routine of the users and involve then in the design process.

This is a diary-based method for the identification of the information demand and acquisition strategies of user representatives during a medium period of time (2 weeks) [9]. Significant in this technique is the systematic annotation of the information demanded by the users in their daily routine and the origin of the information with the current technologies. Especially in the context of smart ecosystems, where a big amount of data is available, having insights on the information demand by users and the origin of this information is essential to improve the development of concrete services in smart ecosystems. Additionally, this technique helps in understanding how people handle big amount of information, e.g. breaking it down in more hierarchical levels or selecting concrete sources for the wished information.

Experience Clip allows users to record their daily interaction with the application with a focus on the emotions and the context in which people are using the mobile application [10]. Context information and emotional state are very relevant aspect, when design for mobile application. This video diary helps designers to understand the daily routine of the users of such systems, their pains and improvement potential.

As we can observe in the table above, all methods try to cover the integration of mobile app in the usage context of the user. Nevertheless, this is just one important aspect for mobile interaction design in smart ecosystem.

This fact calls for the extension of existing methods or the development of completely new ones that support the identified challenges for smart ecosystems.

5 Conclusion

We have shown that mobile-specific development affects classical interaction design through aspects such as context awareness and limited resources. We mapped the ecosystem specifics with regard to mobile interaction design, such as complexity and size as well as standards and regulations that have a significant impact on mobile interaction design.

Overall, none of the considered methods covers every challenge arising from the characteristics of smart ecosystems. However, they are partially useful (see Table 4), which means that it is necessary to tailor these methods or to create a new one in order to address all challenges with regard to smart ecosystems.

As future work, we plan to derive specific requirements according to the whole software development process in consideration with the identified peculiarities of mobile applications and smart ecosystems in order to identify further impacts on mobile interaction design. These will be used to adapt our methods mConcAppt and mPotential and even develop new approaches to take into account the trends and challenges of smart ecosystems. We also call on experts to think in these aspects and work in the development of approaches that address the challenges presented by smart ecosystems.

6 Acknowledgments

This publication is based on the research insights resulting from the Fraunhofer Innovation Cluster "Digital Commercial Vehicle Technology". This cluster is funded by the European Union, the state of Rhineland-Palatinate, and the Fraunhofer-Gesellschaft.

References

[1] Portio Research: Mobile Applications Futures 2013-2017, Analysis and Growth Forecasts for the Worldwide Mobile Applications Market, pp.22-28, 2013

[2] P. Liggesmeyer, “What does an ecosystem have to do with software?”, Fraunhofer IESE Annual Report 2014/2015, Smart Rural Areas, pp.18-19, 2015

[3] J. Knodel; C. Webel: Softwareentwicklung 2020. Business Technology Magazin (www.bt-magazin.de), BT 1.2013

[4] Chen, S. L., Wang, Yue and Tseng, M. M. 2009, “Mass Customization as a Collaborative Engineering Effort”, International Journal of Collaborative Engineering, 1(2), pp. 152-167
[5] H. Muccini, A. Di Francesco, P. Esposito, “Software Testing of Mobile Applications: Challenges and Future Research Directions,” 7th International Workshop on Automation of Software Test (AST), pp.29-35, 2012
[6] S. Hess, F. Kiefer, R. Carbon, “Quality by Construction Through mConcAppt: Towards Using UI-construction as Driver for High Quality Mobile App Engineering,” Conference on the Quality of Information and Communications Technology (QUATIC), pp.313-318, 2012
[7] S. Adam, R. Carbon, S. Hess, F. Kiefer, G. Kutepov, N. Riegel, „Mobility Potential Analysis Method,“ IESE-Report 083.12/E, Kaiserslautern, 2012
[8] M. de Sa, L. Carrico, C. Duarte, “Mobile Interaction Design: Techniques for Early Stage In-Situ Design,” HCI: New Developments, InTech, 2008
[9] T. Sohn, K. A. Li, W. G. Griswold, and J. D. Hollan, “A diary study of mobile information needs,” In Proc. of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08). ACM, New York, pp.433-442, 2008
[10] M. Isomursu, K. Kuutti, and S. Väinämö, “Experience clip: method for user participation and evaluation of mobile concepts,” In Proc. of the 8th Conf. on Participatory design (PDC 04), Vol. 1. ACM, New York, pp.83-92, 2004

Towards Optimizing the Sunburst Visualization for Smart Mobile Devices

Ragaad AlTarawneh[1], Shah Rukh Humayoun[1], Abdel-karim Al-Jaafreh[2]
Computer Graphics and HCI Group
University of Kaiserslautern
Gottlieb-Daimler-Str. 67663 – Kaiserslautern, Germany
[1]{tarawneh,humayoun}@cs.uni-kl.de,
[2]aojaafreh@gmail.com

Abstract. Visualizing large hierarchies is one of the key applications in the information visualization field. The sunburst visualization is one of the space-filling techniques that produce compact representations for large trees via using some of geometry features and solid areas to show the hierarchy. However, the limited screen sizes and interaction styles of current mobile devices bring new challenges for such visualization techniques. Therefore, optimization and customization in the existing techniques are needed in order to make them suitable for the current mobile paradigm. In this paper, we propose some optimization solutions and interaction options for the sunburst visualization in order to increase the viewers' ability in understanding large tree hierarchies and to enable them in interacting more intuitively with the resulting visualization. We also report results of a brief preliminary user study that we conducted with 15 participants.

1 Introduction

Smart mobile devices (i.e., smartphones, tablets, smart watches, etc.) are now an essential part of our daily life activities. The current mobile paradigm is different from the desktop paradigm in many ways due to different factors such as limited screen sizes, new interaction styles (e.g., multi-touch gestures interaction) [3], mobility, context awareness, etc. Many times it is required to represent the data in visual forms in order to give the users the ability to get insight about the data more accurately and efficiently. However, the limited screen sizes of mobile devices and the current mobile interaction styles bring new challenges for visualization techniques. This become more critical when visualizing hierarchical data, as in this case the main concern is to handle the large-seized trees over the limited display sizes.

The Sunburst visualization is a well-known approach in space-filling techniques due to its intuitiveness in conveying the hierarchical behavior of large trees [10]. That's why it has been used in many different domains, e.g., some utilization can be found in [9, 10, 14]. The Sunburst technique could also

be useful to be used on mobile devices due to its compact representations for large trees and presenting the resulting visualization in a circular fashion, which lowers the changing affect of the overall view when rotating the mobile device. However, optimizations and new interaction styles are required in Sunburst technique in order to increase its suitability for the current smart mobile devices.

Targeting this concern, we focus on proposing some optimization solutions and interaction options in the Sunburst (i.e., a *colorization* approach, a *focus+parents-levels-overview* navigation approach, zooming interaction, and a *labeling* scheme) while keeping in mind the limited screen size and the interaction styles of the current mobile devices. These proposed optimization solutions and interaction options would increase the readability of large tree hierarchies and help the mobile users to interact more intuitively and efficiently with such visualizations. We also report a preliminary user study that we conducted to measure users satisfaction level towards our colorization approach.

The reminder of the paper is structured as follows: First, we briefly provide some related work. Then we present our proposed optimization solutions and interactions options, followed by the user study. Finally, we conclude the paper.

2 Related Work

Space-filling techniques can be subdivided in two categories: the *Space-Division* layouts and the *Space-Nested* layouts. In the case of Space-Division layouts, the parent-child relation is depicted implicitly by attaching the children to their parents. The Sunburst algorithm, originally proposed by Stasko and Zhang in [10], uses radial or circular space-filling techniques. A general image in the developer community is that the radial layout methodology conveys a better hierarchy structure without sacrificing the efficient screen-space usage [4, 10]. Whereas in Space-Nested layouts, the child-parent relationship is drawn using nested boxes in which the children are placed within their parent nodes. Few examples can be found in [2, 12, 13].

Yang and Matthew utilized the Sunburst technique to build a tool for supporting the circular distortion to make it more interactive and scalable [14]. In [6], Rodden presented a nice interactive representation to visualize the sequences and the paths that users tend to follow while using the Youtube website.

Recently, researchers have started working in the direction of visualizing data on touch devices and smart mobile devices using different approaches,

e.g.: Sadana and Stasko in [8] provided their approach of visualizing interactive scatterplots on tablet devices; Baur et al. presented TouchWave in [1] for visualizing stacked graphs on touch devices; while Rzeszotarski and Kittur presented TouchViz in [7] for visualizing multivariate data on tablet devices.

3 Optimizing the Sunburst for Mobile Devices

In this section, we provide details of our proposed optimization solutions and interaction options to make Sunburst suitable for the smart mobile devices. As a case study in this paper, we visualize the hierarchical structure of the software system of a robot called RAVON [5] (see Figure 1). The RAVON software system contains four main packages (i.e., *libraries*, *projects*, *tools*, and *others*), where each package contains a group of sub-packages except the "other" package. In total, RAVON has 3152 software files and are divided by three extensions (i.e., ".cpp", ".h", and ".hpp").

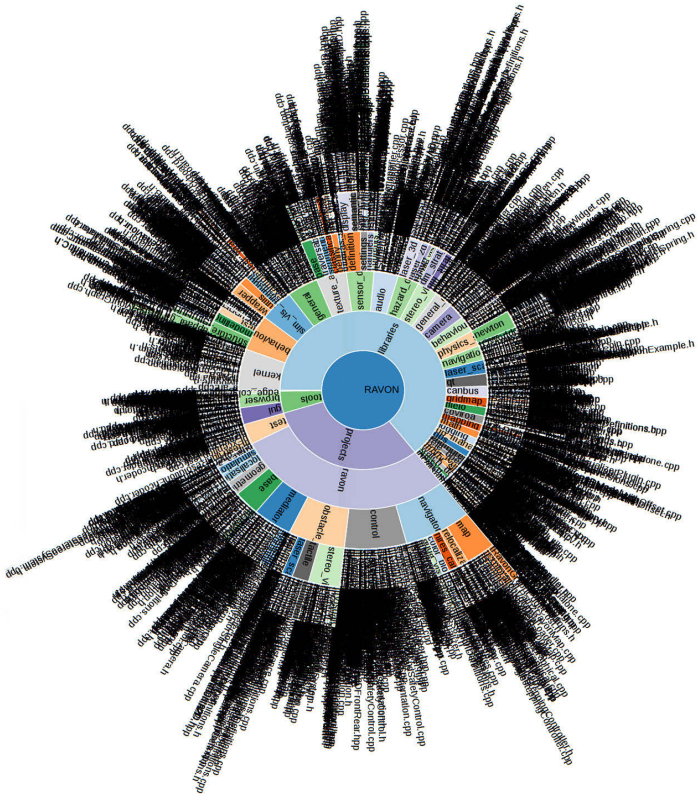


Figure 1. Visualization of the RAVON software system hierarchical structure using the standard Sunburst. It also shows the labels of nodes, which are either packages or class names.

3.1 Enhancement of the Sunburst Colorization Style

Generally, space-filling techniques use solid areas to represent nodes where each node has a specified color. Colorization is the first thing that is noticed by human eyes [11]; therefore, we focus on enhancing the Sunburst visualization using the color dimension strength. For this, we create a new dimension for the visualized node, based on not only the size dimension but also on the color dimension. In this regard, we have two meanings for the colors in our visualization: the *branching* and the *depth*. Each node takes a color based on its parent's color where the brightness of this color depends on this node's depth level.

For example, our case study the RAVON software data structure has four branches (i.e., *libraries*, *projects*, *tools*, and *others*). Therefore, the resulting visualization has four different colors (see Figure 2). These four parent nodes are the darkest nodes in their branch while the children of these parents are lighter in color compared to them and so on until the leaf nodes, which are the lightest nodes. This solution also helps the viewers in keeping the overall structural map in their mind when they browse the lower level of details, as the colors of lower branches indicate their parent branch.

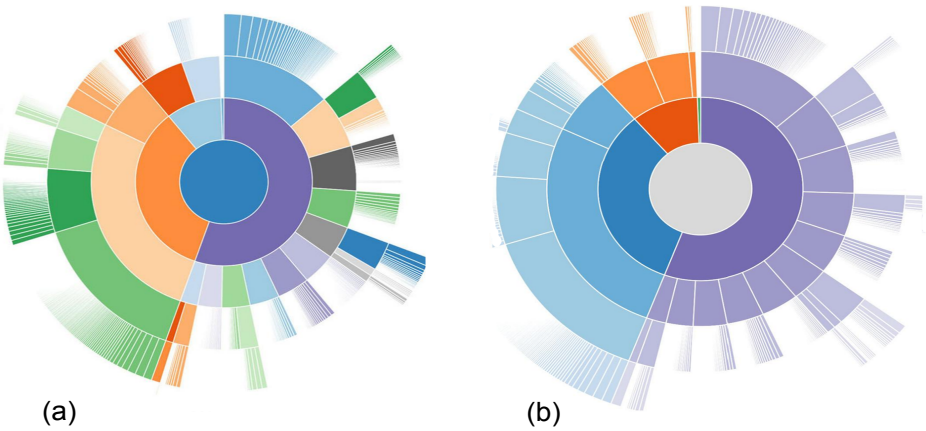


Figure 2. Sunburst before (a) and after (b) the application of our *colorization* approach.

3.2 The Focus+Parents-Levels-Overview Navigation

In order to navigate intuitively through the different levels of details, we propose a navigation style called the *focus+parents-levels-overview* navigation. This navigation scheme combines the *folding* and *unfolding* interactions with a customized form of *focus+context* approach while keeping the requirements of limited screen sizes and the interaction styles of smart mobile devices.

In this regard, when users tap on a particular node, a new view is provided to show its children and lower levels, where the root of this new view is the selected node. However, instead of visualizing the selected node as the root node in the center of the visualization we visualize the selected root and its siblings in a circular fashion inside the new presented view, as shown in Figure 3.b. When users navigate to further lower levels, all the parent nodes and their siblings are visualized one by one in a circular form at the center of the newly created view (see Figure 3.c). In the case if the parent node and its siblings have the same color then the parent node is highlighted in the inner circle (see Figure 3.c).

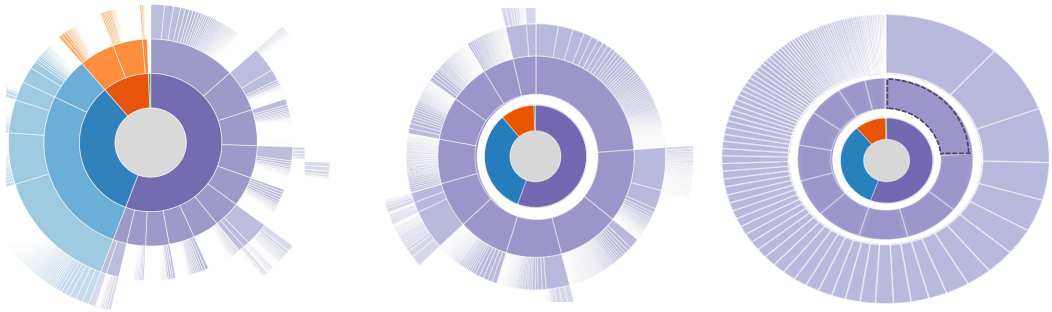


Figure 3. (a) Sunburst using our *colorization* approach showing the overall RAVON software structure. (b) The new view after tapping the "libraries" node, here the inside circle shows the parent node and its siblings. (c) Further navigating the "kernel" node, here are two parent-level inside circles where the "kernel" node is with highlighted border in order to show it as the root of this view.

Taping on any inside parent circle brings the user back to that level of details, where the root is supposed to be the parent of the tapped circle in the newly created view. While double tapping any particular node in an inside parent circle brings the user to the level of details where the double tapped node is supposed to be the root of the newly created view.

This *focus+parents-levels-overview* navigation approach provides several benefits in the context of mobile devices. In desktop environments, normally a separate view is provided to give the overall context so that users do not lose the context when they navigate deep in the hierarchy. Providing a separate overall view is not so feasible in the case of mobile devices. Our approach gives users the important overall context, i.e., all the parents and their siblings, which is enough for having an idea about the overall hierarchy. This also endorses the current mobile paradigm theme, i.e., showing only the relevant and necessary information while hiding the less important information. Also, our *colorization* scheme helps here users in understanding their current level of navigation.

Overall, our proposed navigation approach not only saves the space without affecting much the visualized information, but it also provides intuitive interactions with the resulting visualizations that are more suitable for the current mobile paradigm.

3.3 Zooming

In the cases of large trees (e.g., our RAVON case study), it is sometimes difficult to see low-level details in the produced visualization due to the limited screen sizes of mobile devices. Therefore, it is needed to provide instant distortion and zooming facility so that users can directly explore the hierarchy and other details from the current level of details. In this regard, we provide two options:

In the first option, we provide instant zooming interaction to any view level. This activates when viewers apply long press on any part of the Sunburst visualization, which results in zooming of that particular part. For this, we divide the overall Sunburst circular view into six slices and the corresponding slice of long pressed is then zoomed (see Figure 4.a). The selected area can also be further zoomed by again applying the long press interaction on the previously zoomed area. This time we divide the already zoomed slice into further three slices for providing further zooming. By simple tapping on any particular node in the zoomed area would open the new view where the tapped node would be the root of the new opened view. On the other hand, tapping any place outside the zoomed view would close the zoomed window and would show the current view.

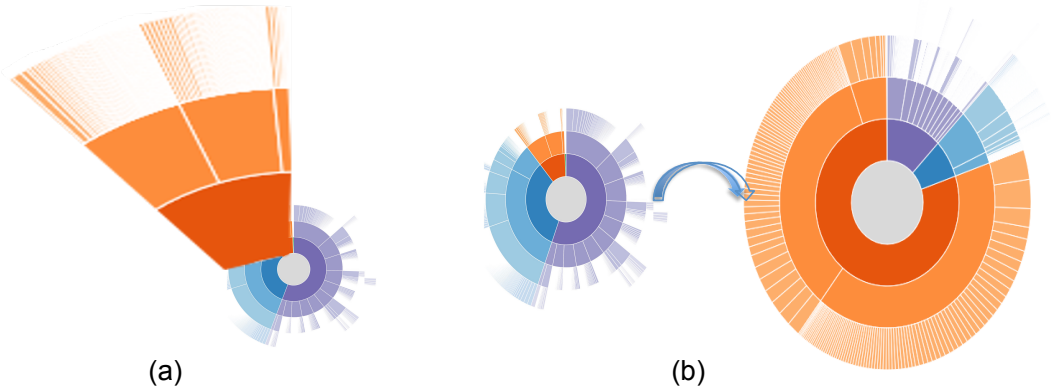


Figure 4. (a) Long pressing on a particular part of Sunburst brings a new view showing the selected slice in zoomed form while the overall Sunburst is scaled down and shown at the bottom side of view. (b) In the second option, long pressing on a particular first-level children node enlarges this node and its further children nodes inside the Sunburst while scaling down other sibling nodes and their children nodes inside the Sunburst.

In the second option, long pressing a particular first-level children node of the current view enlarges this node and its children inside the Sunburst while scaling siblings of the selected node and their children inside the Sunburst. Figure 4.b shows the new view after long pressing the “tools” node (i.e., in orange color) in the original view. This new view scales up the “tool” node and its children while scales down siblings of the “tool” node and their children. Tapping any area outside the enlarge part would bring back the original view.

3.4 Enhancement of the Sunburst Labeling Style

We also focus on optimizing the labeling style in the Sunburst in order to make it suitable for mobile devices. The basic idea is to show the nodes’ names in the visualization without causing extra occlusion. For this, we propose two options:

3.4.1 The External Labeling

In this labeling style, we enhanced the solution proposed by [6] in which labels are visualized using the external SVG elements outside the sunburst. In our case, the top SVG elements (as shown at the top side of Figure 5.a) show the path from the overall root node of the tree to the root node of the current view. However, we enhanced it by showing the labels of all the one-level children nodes of the root node of the current view at the bottom of the view in a scrollable horizontal bar (see Figure 5.a). This option not only helps in reducing the cluttering that might appear in case of large hierarchies, but it also increases the interaction speed with the visualization. Moreover, this light solution offers some natural interactions with the resulting visualization. For example, tapping a label in the bottom horizontal bar highlights the corresponding node in the view while double tapping a label opens a new view making the selected node as the root node of the new view. On the other hand, swiping over the nodes in the view highlights the corresponding labels in the bottom horizontal bar.

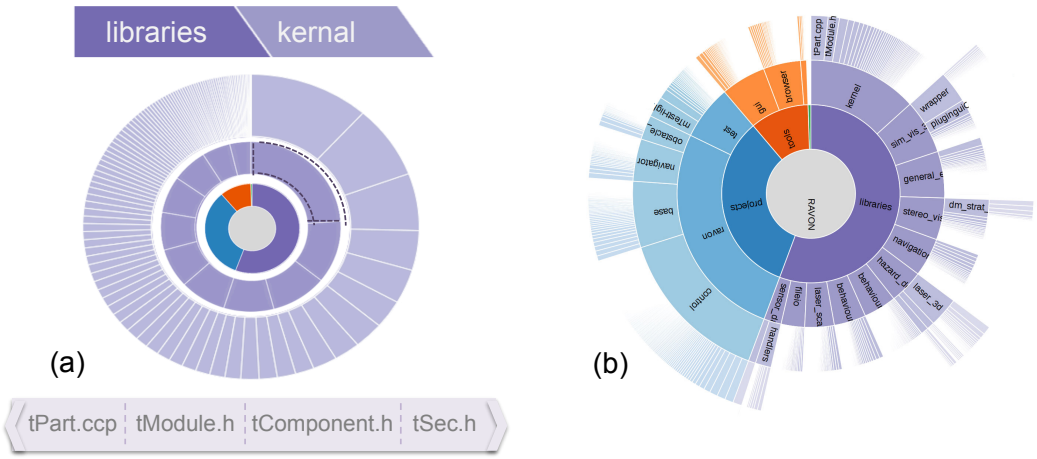


Figure 5. (a) Labels visualizing using the external SVG elements on the top of the Sunburst while the scrollable horizontal bar at the bottom shows names of one-level children nodes. (b) Application of the internal labeling.

3.4.2 The External Labeling

Sometimes it is necessary to show labels inside the nodes. However, using this option with normal node-label length leads to extra cluttering and extra noise in the resulting visualization, as it is shown in Figure 1. Therefore, this solution is not suitable for mobile devices. To optimize this solution, we propose to show only those labels that could fit inside the geometry of their nodes (see Figure 5.b). However, through zooming option, users would be able to see other nodes' labels when they are in the zoomed form, as then these nodes would have more geometrical area.

4 The Preliminary User Evaluation Study

We conducted a preliminary user evaluation study in order to measure participants' satisfaction level towards our colorization scheme. In this study, 15 students participated with an average age of 25 years. Most of these participants were from the software engineering background with expertise in software systems architecture and software metrics, as our case study visualization was based on the RAVON software system. The evaluation study consisted of the following task and goal:

- **Task:** Test the colorization affect.
- **Goal:** The goal of this task was to get the users' feedback with regard to the influence of adding the colorization styles on conveying the hierarchy clearly.
- **Configurations:** The RAVON software system was visualized on a tablet device using two different visualizations: the first one was with

random colors while the second one was with the proposed colorization approach. These both cases are shown in Figure 2.

At the end of the test, we asked the participants to give their feedback through closed-ended questionnaires based on a likert-scale (between 1 and 5 where 1 meant strongly disagree while 5 meant strongly agree and the sixth option was "Don't know").

We asked the participants whether they used the variation in colors to understand the tree structure. In this regard, 12 participants out of 15 participants agreed that they used the variation in colors to detect the hierarchy (see Figure 6). Most of them reported loudly during the test that it is easy to detect the hierarchy levels or branches using the variation in colors. While in the case of using random colors in the provided visualization (see Figure 2.a), participants were almost neutral as they used the containment representation provided by the Sunburst visualization to get the hierarchy rather than using the color cue.

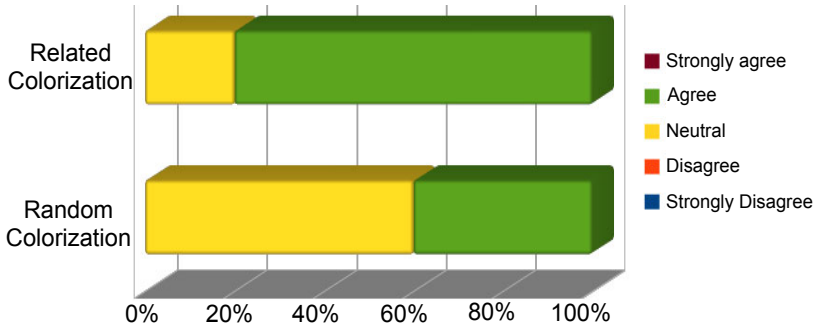


Figure 6. Participants' feedback regarding their preference of using the color cue in finding the tree structure.

5 Concluding Remarks

In this work, we proposed some optimization solutions and interaction options for the sunburst visualization in order to make it more suitable and readable on smart mobile devices. For example, we proposed a *colorization* approach in which each branch has its own color and the levels in a branch have different brightness values of the same color of their parent. This solution was inspired from the work of Tennekes and Jonge in [11]. Using this solution with our *focus+parents-levels-overview* navigation approach provides an intuitive solution for mobile users to understand and navigate large tree hierarchy on the limited screen sizes of mobile devices. Further interaction options were provided to make the interaction more natural according to the current mobile paradigm.

The results of the conducted preliminary user evaluation study showed high users satisfaction towards our *colorization* approach, as they agreed that it helped them in detecting the hierarchy more intuitively. However, this was very limited study with focused only on the *colorization* approach. In the future, we intend to conduct large-scale evaluation studies where we would evaluate all the proposed optimization solutions and interaction options using different data sizes and device sizes in order to check the effectiveness and feasibility of our solutions in general on mobile devices.

References

[1] D. Baur, B. Lee, and S. Carpendale. TouchWave: Kinetic Multi-touch Manipulation for Hierarchical Stacked Graphs. ITS '12, 255–264, 2012.

[2] I. Herman, G. Melanon, and M. S. Marshall. Graph Visualization and Navigation in Information Visualization: A survey. IEEE Trans. on Vis. & Com. Grap., 6(1): 24–43, 2000.

[3] S. R. Humayoun, F. A. Rupprecht, S. Hess, and A. Ebert. Adding Multi-Touch Gesture Interaction in Mobile Web Applications. Human-Computer Interaction, Advanced Interaction Modalities and Techniques Lecture Notes in Computer Science, Volume 8511, 48-57, 2014.

[4] H. H. Kagdi and J. I. Maletic. Onion Graphs for Focus+Context Views of UML Class Giagrams. VISSOFT '07, 80–87, 2007.

[5] M. Proetzsch, T. Luksch, and K. Berns. Development of Complex Robotic Systems Using the Behavior-based Control Architecture iB2C, volume 58. North-Holland Publishing Co., Amsterdam, The Netherlands, The Netherlands, Jan. 2010.

[6] K. Rodden. Applying A Sunburst Visualization to Summarize User Navigation Sequences. Computer Graphics and Applications, IEEE, 34(5), 36–40, September 2014.

[7] J.M. Rzeszotarski and A. Kittur. TouchViz: (Multi)Touching Multivariate Data. CHI '13 Extended Abstracts, 1779–1784, 2013.

[8] Sadana, R. and Stasko, J. Designing and implementing an interactive scatterplot visualization for a tablet computer. AVI '14, 265-272, 2014.

[9] J. Stasko, R. Catrambone, M. Guzdial, and K. McDonald. An Evaluation of Space-Filling Information Visualizations for Depicting Hierarchical Structures. International Journal of Human-Computer Studies, 53(5), 663–694, 2000.

[10] J. Stasko and E. Zhang. Focus+Context Display and Navigation Techniques for Enhancing Radial, Space-Filling Hierarchy Visualizations. Info Vis '00, 57–65, 2000.

[11] M. Tennekes and E. de Jonge. Tree Colors: Color Schemes for Tree-Structured Data. VIS '14, 2014.

[12] J. J. Van Wijk and H. van de Wetering. Cushion Treemaps: Visualization of Hierarchical information. Info Vis ’99, 73– 78, 1999.
[13] C. Weiwei. A Survey on Graph Visualization. Hong Kong University of Science and Technology Clear Water Bay, Kowloon, Hong Kong, 2007.
[14] J. Yang, M. O. Ward, and E. A. Rundensteiner. Interring: An Interactive Tool for Visually Navigating and Manipulating Hierarchical Structures. Info Vis ‘02, 77–84, 2002.

Meanings of a Blurred Mobile-Home Context for People aged 50plus

Thomas Meneweger[1], Marianna Obrist[2], Manfred Tscheligi[1]
[1] Center for Human-Computer Interaction, University of Salzburg, Austria
[2] School of Engineering and Informatics, University of Sussex, UK
firstname.lastname@sbg.ac.at
m.obrist@sussex.ac.uk

Abstract. Due to an increased mobile lifestyle facilitated through the proliferation of mobile and ubiquitous technology, the meaning of home is changing and loses its particularities. In this paper, we explored how 'home' is conceptualized by people aged 50plus, being a fascinating user group when looking at the transition into an area of increased mobility. We conducted in-depth qualitative interviews, demonstrating that people aged 50plus have two distinct viewpoints on home: (1) inside view, understanding home as the spatial inside of their personal living space, and (2) outside view, that extends home beyond the personal living space. We show how these meanings are constructed based on the spatial and social dimensions of home. With this study we aim at providing designers in the mobile application domain a foundation for reflecting on the home context, opening up new opportunities for intuitive applications bridging the domestic and mobile life.

1 Introduction and Motivation

The increased penetration of people's everyday life by mobile and ubiquitous systems and services contributes to a blurred understanding of what the home is and how the home and the mobile context are interrelated [7], [10]. Mobile applications that promote wellbeing and allow for social connectedness are especially beneficial for people aged above 50 [5]. We believe that our research can serve as a theoretical basis for reflecting and discussing on future designs of intuitive mobile applications for people above 50 within an increasingly mobile era. By revealing the mental models of home for people aged 50plus, we aim at providing theoretical grounding for interaction designers and researchers to explore intuitive interaction concepts for the blurred intersection of the home and the mobile context as an interrelated design space.

Related work explicitly addressing the blurred intersection between the home and mobile context is rather rare. Only a few studies are concerned with the domestic area, which is increasingly interwoven with mobile [6] and ubiquitous technologies [10]. Petersen et al. [7] made a first step towards reaching a richer understanding of a blurred mobile-home context by

investigating peoples practices and technology usage to make them feel homey in other places. They identified different notions of feeling at home in other places. Further research endeavors dealing with the intersection of the home and mobile context are exploring work-home boundaries and how to maintain them (e.g., [2]).

With our work, we want to emphasize the need to explore the basic understanding how people aged 50plus experience a blurred mobile-home context (an extension of home due to increased mobility).

2 The Study Setup

To explore subjective meanings of the home for people aged 50plus, we conducted open-structured qualitative interviews in a Western European country and applied the documentary method [1] to analyze the qualitative data.

Our main research question was: *What do people aged above 50 years mean when talking about the home?* We set up the interview guideline based in the idea of the narrative interview developed by Schütze [8]. The interviews were conducted with six female and two male participants. The participants were aged between 52 and 67, with a mean age of 60 years. The interviews were audio-recorded and transcribed for further analysis.

3 Findings

3.1 Meanings of Home

A key aspect of the emerged *meanings of the home*, based upon the spontaneous descriptions collected from the participants, is the distinction between the *inside* and *outside views* on the home.

3.1.1 Inside View on the Home

Participants with an *inside view on the home* talk basically about the interior of their house or flat. The door of their house/flat represents the external border of their home. Their descriptions about the home are focusing on what is happening inside their house/flat as a physical space. For example, a participant pointed out that for him/her, the home is *"That you can close the door"* (P7).

Within this *inside view* on the home two subtypes were identified. Participant's expressions subsumed within subtype (a) highlight that home is a

personal area for them (*"Here, my flat, the private sphere"* (P1)). Apart from privacy issues, participants' perceptions of the home are associated with the home as a refuge from society, a kind of a haven to escape social life and forget about what is happing outside. Within subtype (b) of the *inside view* on the home, participants express a strong emotional attachment to their home. Participants mention things such as comfort, coziness and feeling of security to describe their meaning of the home, e.g., *Where you feel comfortable, it is simply my kingdom, it is my realm, where I live how I like."* (P5).

3.1.2 Outside View on the Home

People with an *outside view* on the home focused within their descriptions of the home on the surrounding area of their house/flat. They describe mainly what is happening outside their house/flat. Within this type the understanding of the home moves beyond the physical borders of a house or a flat, which the following quote on the description of the home illustrates: *"Well, at first* [name of the city anonymized]. *Then it means for me to live in the city center"* (P8).

Within the *outside view* on the home two subtypes were identified. Subtype (a) encompasses descriptions of the spatial environment of the house/flat, which focuses on the places surrounding the house/flat (e.g., the village, the city). One participant points out that he only has the feeling to have left his home when *"passing the railway bridge of the city"* (P8). The other way round, this participant has the feeling to come home when coming back to the city over the railway bridge. Subtype (b) within the *outside view* focuses on the social environment of the home. Participants relate to people living nearby their house/flat, such as their neighbors and friends. Participants' mapped into this subtype have a strong emphasize on the social character of the home and feeling connected with other people, e.g., for *"being together"* (P4).

3.2 Strategies of Extending the Home

Beyond the two main types of *meanings of the home*, we additionally identified two different *strategies of extending the home*. Within the participants' understanding of the home, the spatial dimension of the home often lacks congruency with the social dimension of the home. This is the case when, for example, family members, friends and relatives are living outside the physical border that participants have previously described as their home. However, these people are still considered as being part of his/her home. We found that participants adjusted their meaning of the home during the interview, according to two strategies on top of the initially emerged meaning

of the home. On the one hand, by "extending" the home spatially to external places, on the other hand, by "taking" people (from places outside the "home") inside their homes.

3.2.1 Taking People Inside from the Outside

During the interview this strategy was applied by participants who outlined that people, who are initially (e.g., at the beginning of the interview) located outside the space that was prior defined as the home, and then incorporated into the space inside the house/flat. For example, participants that were assigned to this strategy mentioned that specific people like family members (e.g., children living in a distance), friends and relatives are part of their subjective definition of the home. However, only when these people are visiting them in their houses/flats. Some participants mentioned that their house/flat means home to their children who are not living anymore in their parents house/flat (P4, 5 and 6). This is of course what the interviewees think, we do not know if their children experience it in a similar way. This strategy corresponds to the *inside view* on home[1].

3.2.2 Extending the Home to Outside Places

The second strategy of extending the home relates to the extension towards outside places, which even go beyond the *outside view* on the home previously defined. Participants applying this strategy extended their understanding of home to remote places by mentioning, for example, that they feel at home when they are staying with their children or when they are visiting friends and relatives (P3 and 4). Other participants even referred to remote places such as a friend's house, that this place means home for them (P2 and 8). For example, P8 referred to her office to be home for her as well. This person somehow "disarranges" the typical structure of her home (kitchen, living room, bedroom, etc.), by saying that the office is a living room for her (even more than the "real" living room in her flat). The strategy of *extending the home to outside places* can be identified in the descriptions of the same people fitting into the *outside view* on the home.

---

[1] For P1 (showing an *inside view*) we couldn't identify any strategy, because this person didn't refer to any social dimension of home within his descriptions.

4 Reflections on a Blurred Mobile-Home Context

Our findings indicate that people aged 50plus are impacted by an increased mobile lifestyle. The social dimension of these people's meaning of home is obviously influenced by mobile developments, as the social dimension is in most cases extended beyond the spatial dimension of home. Due to an increased proliferation and penetration of mobile and ubiquitous technologies people can easily get connected with family members (e.g., children) and friends, who are living at a distance. For example, people with an *inside view* driven understanding of home (referring to the physical space inside their house/flat) are regardless their affinity for mobile technologies still affected by mobile developments, as their children and friends might use mobile technologies to frequently stay in contact with them from the outside.

The two main types of *meanings of the home* indicate that the understanding of home can be split into two groups. A rather "traditional" group referring to the home as bound to the physical space of a house or a flat (*inside view*), which has been extensively investigated by "traditional" studies [3], [9]. However, another more "progressive" group referred to an extended understanding of the spatial dimension of the home, beyond the physical space of the house (*outside view*), which received only limited research interest up to now. Petersen et al. [7] and Lyngaard et al. [6] have investigated this blurring between the home and mobile context for a "progressive" and young user group [7].

In contrast to that, previous work focusing on the younger generation, we focused on people aged 50plus. Understanding how this user group experiences and conceptualizes home under the changing circumstances helps us in HCI research to establish a strong foundation for improving the design of applications for the mobile-home context, as it allows a systematic reflection on the transition and associated meanings and experiences. This is similarly valuable as recent research studying the transition from manual to autonomous driving and the changing meaning of a car and driving [4]. Designers will need to carefully study such findings, as they will highly influence the design process of new interfaces and the creation of pleasurable and satisfying experiences.

Our findings indicate that our participants actively constitute the meaning of home, comprising the social elements (i.e., children, relatives, friends) and mentally (but actively) shape them towards a specific spatial arrangement. The revealed *strategies of extending the home* represent two different ways to align the social elements of the home with the spatially dominated concept of the home, and therefore shaping the understanding of the mobile-home context.

Thus, our findings underline that home is a complex concept that is not just simply there and as well influenced by mobile applications. Future research in the blurred mobile-home field as well as interaction designers should strive for exploring the interrelations and dynamics of the social and spatial dimensions of the meanings of home with respect to people's mobility.

5 Conclusion

With the presented insights and reflections we aim to provide a theoretical basis for researchers as well as interaction designers in HCI for future explorations of mobile applications. More specifically, we found different conceptualizations of a blurred mobile-home context for people aged 50plus that might influence how people interact with mobile devices and applications within their everyday environments. How to incorporate these conceptualizations in the design of interactions and visualizations of mobile applications that aim to be intuitive will be both a challenge and promising opportunity for future developments.

Acknowledgements

The financial support by the Austrian Federal Ministry of Science, Research and Economy and the National Foundation for Research, Technology and Development is gratefully acknowledged (Christian Doppler Laboratory for "Contextual Interfaces").

References

[1] Bohnsack, R. Documentary Method and Group Discussions. In Qualitative Analysis and documentary Method in International Educational Research. R. Bohnsack, N. Pfaff, and V. Weller, Eds. Budrich (2010), 99-124.

[2] Cox, A.L., Bird, J., Mauthner, N., Dray, S., Peters, A., and Collins, E. Socio-technical practices and work-home boundaries. In Proc. of MobileHCI'14 (2014), 581-584.

[3] Crabtree, A., Mortier, R., Rodden, T., and Tolmie, P. Unremarkable Networking: The home network as a part of everyday life. In Proc. DIS'12 (2012), 554-563.

[4] Lee, K.J., Joo, Y.K., and Nass, C. Partially intelligent automobiles and driving experience at the moment of system transition. In Proc. CHI'14 (2014), 3631-3634.

[5] Lindley, S.E., Harper, R., and Sellen, A. Desiring to be in touch in a changing communications landscape: attitudes of older adults. In Proc. CHI'09 (2009), 1693-1702.
[6] Lynggaard, A.B., Petersen, M.G., Gude, R. and Mortensen, M. Home awareness: connecting people sensuously to places. In Proc. DIS'10 (2010), 416-418.
[7] Petersen, M.G., Lynggaard, A.B., Krogh, P.G. and Winther, I.W. Tactics for homing in mobile life: a fieldwalk study of extremely mobile people. In Proc. MobileHCI'10 (2010), 265-274.
[8] Schütze, F. Biographieforschung und narratives Interview. In Neue Praxis 13 (1983), 283-293.
[9] Venkatesh, A., Kruse, E. and Shih, E.C. The networked home: An analysis of current developments and future trends. In Cognition Technology and Work, 5, 1 (2003), 23-32.
[10] Wyche, S., Sengers, P., and Grinter, R. E. Historical analysis: using the past to design the future. In Proc. Ubicomp'06 (2006), 35-51.

Workshop:
First International Workshop on Interactivity in Healthcare Systems (IHS 2015)

Context Awareness in Communication around Fall Handling with PERS

Jan Van den Bergh[1], Shirley Elprama[2], Jasmien Decancq[2], An Jacobs[2], Karin Coninx[1]

[1]Hasselt University - EDM - iMinds
Wetenschapspark 2
3590 Diepenbeek, Belgium
firstname.lastname@uhasselt.be

[2]VUB - SMIT - iMinds
Pleinlaan 9
1050 Brussels, Belgium
firstname.lastname@vub.ac.be

Abstract. Assuring timely intervention after falls is important to enable older adults to live independently for a longer time. There are two strategies where technology could assist timely intervention: 1) automated fall detection and 2) handling of falls - the process of sending help to a fall victim - using a personal emergency response system (PERS). This paper presents first insights on using sensors not only on the patient's side but also on the caregiver side. We present the results of two studies that were part of a more encompassing approach to improve fall handling and draw conclusions based on these results. The first study investigated the willingness of informal caregivers to (automatically) share information that can influence timely fall handling. Based on the results, a prototype of a fall handling system was made. The second study evaluated the prototype with the different types of users that would use this system. Results indicate that while the envisioned users were in general open to context-aware fall handling, fall-back scenarios need to be present because the needed technology may not be available to all users at the same time and some people might not be able or willing to use the technology.

1 Introduction

Falls among older adults can cause major physical and psychological harm and, they also are a major cost to the people affected as well as society [1]. Therefore, a great amount of research is being conducted on fall prevention and fall detection. However, there is also a need to further investigate fall handling, which we define as the process of sending help to a fall victim.

Fleming and Brayne [2] report that many people having a fall alarm did not use it when they fell. A PERS (personal emergency response system) is generally used to indicate an alarm system that contacts a call center for help when a need is detected by a push on a button or using connected sensors. Several reasons can be identified for non-use of PERS: not able to reach the button, not developing a habit to wear it and not seeing its advantage [3]. Many others did not perceive a need to have a fall alarm themselves.

Heinbucher et al. [4] came to a similar observation [4] and further noted that the notion of relevance, rather than satisfaction was an important predictor of PERS use.

While several fall detection services have been proposed that include notification to social contacts [5][6][7], the focus was mainly on the fall detection. For most (if not all) of these no evaluation of the fall handling system is provided. These systems use various approaches to contact caregivers or a call center, but rely on phone calls and/or text messages.

We thus identify some opportunities to improve fall handling: **Usage**: convincing care receivers to use the system (when they are in need); **Caregiver search**: identify the most useful person to attend a fall; **Information exchange**: inform caregivers when needed.

In this paper, we build on preliminary results on using context-awareness in this situation in call center interfaces that were presented by Van den Bergh et al. [16]. We extend this work through two studies we performed as part of a more encompassing approach to investigate the acceptance and feasibility of a context-aware fall handling procedure as part of a PERS system (operated by nurses). These studies provide insights from (potential future) care receivers, caregivers and nurses involved in the PERS system.

2 Related Work

Several studies have explored the willingness of people to share location information with social contacts in non-care contexts. Lederer et al. [8] as well as Consolvo et al. [9] came to the conclusion that the most important determinants to decide whether to share information were who was asking it and why. There seems a common theme that turning off automatic context sharing seems more related to concerns for their peers (modesty) rather than personal concerns (secrecy) in several circumstances [10][11]. A long-term study [12] on surveillance at home showed that while surveillance did not cause higher stress levels, some types of sensors such as cameras, audio and PC surveillance did cause frustrations and even anger, and altered behavior. After a couple of months most people became accustomed to it. Bentley et al. [13] found that adults (ages 21-52) were willing to share travel time with family and friends in a chat app during a field trial lasting 21 days.

Study 1 in this paper contributes to this research as it addresses location and availability sharing from a very specific and different context: sharing in function of a commitment helping a care receiver after a fall.

iFall [5], Living++ [14] and SEHMS [6] are smartphone based fall detection and handling systems that notify the care receiver of a detected fall.

When the care receiver does not cancel the alarm, predefined caregivers get a text message or e-mail. If they call back, the call is automatically answered after a notification sound. There is no PERS integration in this case. Teroso et al. [7] proposed a wearable sensor connected to a smartphone to do fall detection. When a fall is detected, the care receiver gets the options to send a text message (with the GPS position) or call a predefined caregiver. In case of a timeout a text message is sent automatically. The system also has a server allowing caregivers to monitor and configure the system for the care receiver.

These approaches have not been tested with envisioned end users and seem to focus on demonstrating technical feasibility. The emphasis of this paper is on the opportunities and challenges identified with potential future end users of such a system. Our research focused on fall handling rather than fall detection.

3 Study 1: Interviews with Informal Caregivers

3.1 Method

Sixteen interviews with informal caregivers (12F; 4M) were conducted and analyzed based on the following criteria: the aging care receiver had to live independently and not with the interviewed informal caregiver, since a PERS and sharing contextual information seemed less relevant when care receivers live with their informal caregivers.. Their average age was 49.5 years (s.d. 14.2). The corresponding care receivers were on average 82.3 years old (s.d. 6.1).

2nd year bachelor students of communication science conducted and transcribed the interviews with people labeling themselves as informal caregivers (partners, children, other relatives and friends of care receivers) and were selected with purposeful quota sampling. The students received training, a list with interview questions and a movie with a voice-over explaining the contacting of informal caregivers based on location data and availability. Examples of interview questions are: Are you willing to share data about your location/digital calendar with the call center in case a call is placed via a PERS of your care receiver? How detailed may this information be? Who would you allow to have access to this information?

3.2 Results and Discussion

The interviews were iteratively coded by one of the authors using grounded theory [15] starting from open coding, axial and then selective coding.

Dedoose, a web-based application for qualitative research, was used. After coding 16 interviews, data saturation was reached; no new information appeared.

In general, the informal caregivers had a more positive attitude towards sharing information with the call center about their location than sharing information about their availability. This was mainly due to privacy concerns. Based on the results of study 1, we can state that caregivers are prepared to share their location with the call center and even with other caregivers. Several respondents even wanted to share the location with the care receiver. The following quote illustrates this: "*It is for example possible that I go to the store one day and that I'm further away than my brothers. (...) Then it makes sense that the call center calls my brother first, because he would be here first to help my mother. This is a very big solution for me in case I would not be present in the house*." (F, 49). A few caregivers did not see the value for the care receiver as illustrated by the statement: "*I don't think this is a good idea, I immediately have the feeling that 'Big Brother' wants to keep an eye on me*" (F, 54). We thus propose to limit sharing location with the care receiver to cases where there is a clear benefit; such as when the care receiver is waiting for a caregiver.

During the interviews sharing availability based on calendar data appeared to be misunderstood as sharing all data contained in the calendar. The following quote illustrates this: "*Well, honestly I do not think that is a good idea*. [Sharing] *my location is one thing, but on top of that they don't need to know what I'm doing there*..." (M, 20). Another concern for using the calendar was that it would be the (only) way to determine who should attend a fall; "*So, you are forced to register* [your activities], *because otherwise the system cannot function, in that way it brings you extra duties (...) and if you don't register, it is possible that they count on you being the first one to be there* [at the care receiver], *while it might be the case that you are sick and your sister is just at home*." (M, 49).

Limited usage and willingness to start using a digital calendar (and maintaining it) excludes it from being the only way to determine availability. Another easy way to specify availability should thus be present.

The majority wanted to receive call status information through message. "*Yes* [I would like to receive a message], *it would be comforting to know what exactly is going on so I can anticipate if such an incident occurs again*." (F, 48). This view was not universal: *No,* [I don't expect a message from the call center], *because my brother will contact me if it is serious*." (F, 54).

Although informal caregivers preferred to be contacted by phone, some were willing to be informed by text messages. We decided to further explore

usage of text messages as it can reduce time spent communicating with caregivers.

4 System Description

The proof-of-concept system consisted of a native Android app for three types of usage that applied most of the design guidelines of Massimi et al. [17]:

A *call centre app* (designed for tablet, Figure 1, left) with a focus on handling a reported fall including the possibility to view information about the care receivers and their context (information from sensors installed at the care receiver's home) as well as information about registered caregivers, including name, contact information, availability and when relevant, location. The app can be used to call a care receiver and send (textual) messages to care givers using instant messaging over XMPP.

A *care receiver app* (designed for smartphones, Figure 1, middle) with the possibility to ask for help via the call center when this is needed (for instance, after a fall). A care receiver can also send her mood using happy / unhappy smileys or contact available caregivers in "less urgent" situations.

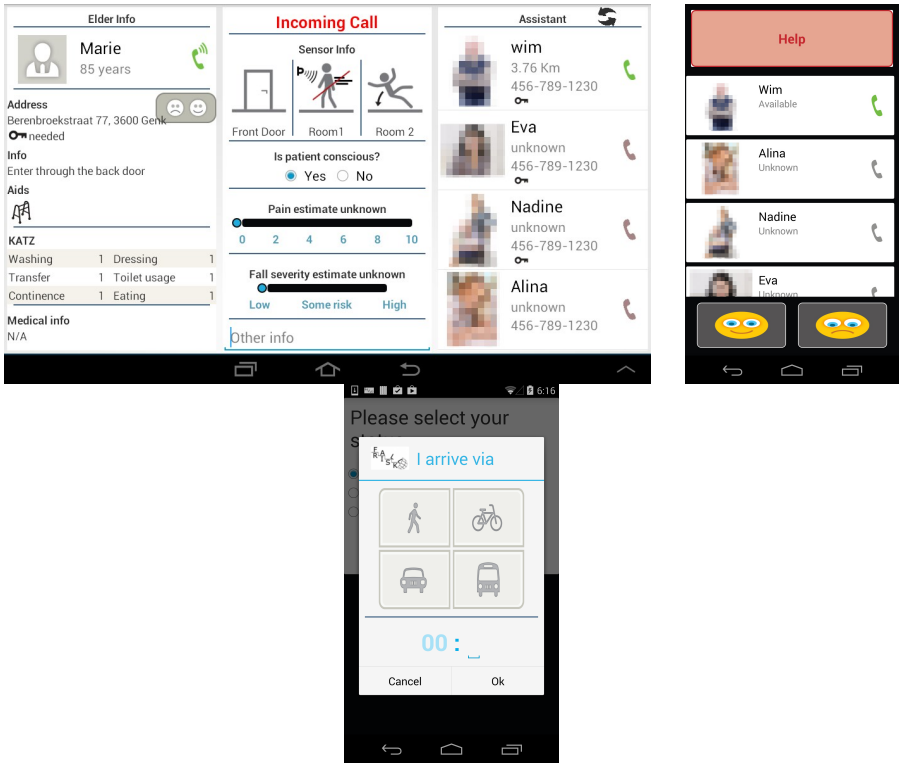


Figure 1. Call centre interface during call (left), main care receiver user interface (middle), a preliminary caregiver screen to specify time to arrival (right).

A *caregiver app* (designed for smartphones, Figure 1, right) with the possibility for caregivers to specify their availability and to receive messages or calls when their care receiver needs help. The app can also automatically detect the location of the caregivers and assist in determining the distance to the care receiver when a fall has been detected.

The app displays one's presence through a colored phone icon. For the care receiver, a green icon is shown when the caregiver is available, when the caregiver is busy or unavailable, the caregiver is removed from view, and when she is disconnected or no information is available, a grey icon is shown. A similar convention is used for the call center although in this case a busy status is also mapped to a green icon. This way caregivers can make a distinction in their availability for call center and care receiver. Voice communication between call center and care receiver apps was simulated using icons. No real voice communication was implemented as the system was designed for evaluation and demonstration of the concept within a single room. The app supports location detection through the Android location application programming interface (API) and availability was determined based on the Android Calendar Provider API. Different calendars can be combined for a single user. One can configure whether presence of a calendar items should reflect "busy" or "unavailable".

5 Study 2: User Tests with Older Adults and Nurses

The goals of Study 2 were (1) to explore the opportunities and challenges of fall handling and (2) to validate the findings of study 1 as well as our design choices.

5.1 Method

Eight user tests following the same protocol were organized. A total of 12 participants took part in the user tests: 8 older adults, including three former nurses (3M; 5F) with an average age of 69.5 (4.5 s.d.) years old and four home care nurses (2M; 2F) with an average age of 41.3 (s.d. 9.0) years old having regular experience with a PERS system. The 8 older adults were purposefully chosen, since they are potential future users.

In four of the user tests two older adults participated and in the next four user tests a home care nurse contributed (none of them participated in Study 1).

Each user test started with *role playing two scenarios*: (1) to experience how fall handling with PERS works and (2) to gain understanding how the

new apps affected the new and proposed fall handling procedure. The scenarios were based on current and envisioned practices of fall handling at a home care organization using information from observations and interviews with employees of this home care organization. The scenarios described what happened after a fall victim needed help. Next, the researchers asked questions to explore the opportunities and challenges of the designed apps and assignments. Questions asked include: Can you describe to us what you see? What do you think it means? An example assignment is: Please indicate your availability on the caregiver app.

In the scenarios, the nurses acted as a nurse in the call center or a nurse responding to an emergency call. In the sessions with two older adults, one played an informal caregiver, while the other played a care receiver. During each session, the apps were shown on a smartphone (caregiver and care receiver app) or a tablet (call center app). Screenshots of smartphone apps were printed in color, so both participants could see the interface during the discussion.

5.2 Results

The interviews were transcribed by students. The results were analyzed similarly as in Study 1, however, different software was used (MS Word for open coding and MS Excel for axial and selective coding) after which the results were written.

During the user tests, older adults quickly accepted their role as care receivers or informal caregivers. Role playing scenarios with the proof-of-concept apps turned out to be a successful method to demonstrate the envisioned use of the apps; we succeeded to let the participants ask questions and share their opinions and concerns regarding the envisioned use of these fall handling apps.

Overall, the older adults were positive about the care receiver app, even though most of them were visibly not used to having a smartphone and were concerned about care receivers' ability to use smartphones. *"I thought, it would be very complicated, but when you see it like this, I think it's really easy for the patient."* (Older adult, F, 72). The same older adult mentioned she liked seeing the pictures and the contact details of the caregivers on the app (Figure) as it was a good reminder of who is who. The older adults viewed the app as more than a fall handling system. They suggested, for example, to use it for communication in general, not only for emergency situations.

Some older adults found it important for a care receiver to view the status of their caregiver. Available and not available were interpreted as was

intended (being or not being available to be contacted to help a care receiver in case of emergency). Both nurses and older adults liked the possibility of the care receiver app to send happy or unhappy smileys. Sending smileys was perceived as an accessible way for care receivers to let their caregivers know how they feel. A couple older adults viewed an unhappy smiley as a trigger to make a phone call to care receivers and one older adult (F, 72) could see herself using it.

Sending the estimated time it would take to get there to the most available emergency contact was valued by the nurses. Multiple older adults stressed that this would make care receivers feel more at ease. In one session, the older adults (who were former nurses) were concerned that care receivers would undesirably contact caregivers also for non-urgent reasons, such as loneliness.

In general, the older adults preferred receiving a text message with a request to help a care receiver to receiving a phone call, as a message was perceived as quick and efficient. An appropriately adjusted volume and ringtone was perceived by the older adults to be a necessity to hear the messages.

In one session, older adults argued a message to all caregivers is always desired in case a care receiver needs help, because even if you cannot help, you will at least know that the care receiver had fallen when you visit him/her the next time. Since nurses handling the falls usually do not know the care receiver, they need additional information*: "Sometimes this* (information) *is on file; we have a file with every patient. Or sometimes I find it on my tablet. Or sometimes, when the patient is conscious, I just ask if there is somewhere a telephone number, can I reach someone or should I call someone?"* (nurse, F, 33).

Most older adults did not mind the automatically calculated estimated duration of travel, but emphasized that one needs to be able to adjust this estimation, for instance in case of a traffic jam. One older adult (M, 65) suggested that it was important to be able to update the estimated duration of travel.

Some older adults considered filling out a report with a few questions, after they 'helped' their care receiver (as part of the scenario) a task for the call center or more knowledgeable people (like a doctor), while others said they would feel more comfortable filling in the report via the phone (and calling the call center). The nurses appreciated the possibility to (get reminders to) fill out the report, because they sometimes forget to write it. *"You know, this* (smartphone)*, you always have it with you. I work full time, and when I come back from a visit, someone has fallen, and then I'm just not in the mood to do this work at home. And then I sometimes forget it."* (nurse, F, 33). They made

suggestions to include specific information with dedicated user interface controls such as sliders.

The older adults seemed to prefer to indicate availability manually, as none of them used a digital calendar. Nurses saw integrating their work calendar as an opportunity. Some older adults linked usage of a digital calendar to sharing caring arrangements with respect to availability to care receiver with call center or sharing planned unavailability. While the nurses were concerned that the caregivers would forget to fill in their availability (as is currently the case with PERS), the older adults perceived providing correct information about their availability as an obligation that comes with being an emergency contact.

The nurses liked the more flexible list of caregivers and the three levels of availability (available, busy, unavailable) seemed handy to receive information about a call, but still being able to finish the work they were doing. *"It could be when working in the day shift* (...) *that I'm doing my round. If there is a call, we should respond of course, we could indicate that we will go, but you can't just drop your other patient."* (nurse, M, 40). The automatic registration and digital approach is seen as useful support, although one nurse expressed concerns about the disappearance of some of the verbal communication in the current system.

In general, the nurses were positive about the call center app: it is easy to read and it gives a good overview of relevant information presented in the app. One nurse insisted to still have the possibility to have verbal communication with both patients and informal caregivers in a crisis, mainly for reassurance. He feared that with digital communication there could be more misunderstandings.

The new approach of availability and the more flexible list of caregivers in the call center app were appreciated by the nurses. Integration of the agenda of patients was suggested. A potential challenge in the app is explaining why the second person in the list was contacted to avoid discussions afterwards.

6 Lessons learned

We could identify and confirm *opportunities* of context-awareness in PERS systems. Relative location, which caregivers are willing to provide, was found to be valuable information during several stages of a fall handling. Caregivers are willing to share availability and it proved useful to facilitate communication.

Several caregivers appreciated that (textual) messages allowed to access information about the care receiver and the fall when they needed it, although they should not be sent in the same manner as SMS, as in earlier work. The

potential speed of using messages was also considered important by informal caregivers. Furthermore, textual messages allow contacting or informing multiple people at once, automated inclusion of relevant information and consultation at a later time. Emoticons showed potential to trigger contact with a care receiver and to lower the threshold of contacting someone, a problem seen with PERS usage.

We discovered *challenges* to using context-awareness. While digital agendas have potential to automatically determine availability, openness to using (and updating) a digital agenda is limited. Several nurses expressed a clear preference to use voice calls to exchange information with caregivers. Use of the advanced features of smartphones by any of the users should not be required. Decisions based on context information need to be traceable to answer questions of caregivers after a fall has been handled.

While the presented studies provide initial insights on how to augment a PERS system with context-aware information in a way that is acceptable to all users of the system, actual behavior can differ from stated behavior. The observed opportunities and challenges should be validated in a larger long-term study. Current PERS users were not included in this study, future research should investigate their attitudes towards the described fall handling system.

Acknowledgements

The iMinds FallRisk project is co-funded by iMinds, a research institute founded by the Flemish Government. Companies and organizations involved in the project are COMmeto, Televic Healthcare, TP Vision, Verhaert and Wit-Gele Kruis Limburg, with project support of IWT. Selim Salman contributed to the realization of the apps.

References

[1] J. Close, M. Ellis, R. Hooper, E. Glucksman, S. Jackson, and C. Swift, “Prevention of falls in the elderly trial (profet): a randomised controlled trial,” The Lancet, vol. 353, no. 9147, pp. 93–97, 1999.

[2] J. Fleming and C. Brayne, “Inability to get up after falling, subsequent time on floor, and summoning help: prospective cohort study in people over 90,” Bmj, vol. 337, 2008.

[3] V. Hessels, G. S. Le Prell, and W. C. Mann, “Advances in personal emergency response and detection systems,” Assistive Technology, vol. 23, no. 3, pp. 152–161, 2011.

[4] B. Heinbuchner, M. Hautzinger, C. Becker, and K. Pfeiffer, "Satisfac-¨ tion and use of personal emergency response systems," Zeitschrift fur¨ Gerontologie und Geriatrie, vol. 43, no. 4, pp. 219–223, 2010.
[5] F. Sposaro and G. Tyson, "ifall: an android application for fall monitoring and response," in Proc. of EMBC 2009. IEEE, 2009, pp. 6119–6122.
[6] J.-V. Lee, Y.-D. Chuah, and K. T. Chieng, "Smart elderly home monitoring system with an android phone," Int. J. Smart Home, vol. 7, pp. 17–32, 2013.
[7] M. Terroso, R. Freitas, J. Gabriel, A. T. Marques, and R. Simoes, "Active assistance for senior healthcare: A wearable system for fall detection," in Proc. of CISTI. IEEE, 2013, pp. 1–6.
[8] S. Lederer, J. Mankoff, and A. K. Dey, "Who wants to know what when? privacy preference determinants in ubiquitous computing," in CHI'03 extended abstracts. ACM, 2003, pp. 724–725.
[9] S. Consolvo, I. E. Smith, T. Matthews, A. LaMarca, J. Tabert, and P. Powledge, "Location disclosure to social relations: why, when, & what people want to share," in Proc. of CHI 2005. ACM, 2005, pp. 81–90.
[10] L. Barkhuus, B. Brown, M. Bell, S. Sherwood, M. Hall, and M. Chalmers, "From awareness to repartee: sharing location within social groups," in Proc. of CHI 2008. ACM, 2008, pp. 497–506.
[11] L. Barkhuus, "The mismeasurement of privacy: using contextual integrity to reconsider privacy in hci," in Proc. of CHI 2012. ACM, 2012, pp. 367–376.
[12] Oulasvirta, A. P., J. Perkio, D. Ray, T. Vahakangas, T. Hasu, N. Vainio, and P. Myllymaki, "Long-term effects of ubiquitous surveillance in the home," in Proc. of UbiComp 2012. ACM, 2012, pp. 41–50.
[13] Bentley, F. R., Chen, Y. Y., & Holz, C. "Reducing the Stress of Coordination: Sharing Travel Time Information Between Contacts on Mobile Phones." in Proc. of CHI 2015. ACM, 2015, pp. 967-970.
[14] M. U. Iqbal, N. Fet, S. Wagner, M. Handte, and P. J. Marron, "Living++:´ a platform for assisted living applications," in UbiComp 2013 adjunct publication. ACM, 2013, pp. 853–860.
[15] Strauss, Anselm, and Corbin, J.M. Basics of qualitative research: Grounded theory procedures and techniques. Sage Publications, Inc, 1990.
[16] J. Van den Bergh, K. Luyten, B. Aendekerk, S. Elprama, A. Jacobs, and F. De Backere, "Interdisciplinary design of a pervasive fall handling system: a case study," in Proc. of PH 2014. ICST, 2014, pp. 235–238.
[17] M. Massimi, R. M. Baecker, and M. Wu, "Using participatory activities with seniors to critique, build, and evaluate mobile phones," in Proc. Of ASSETS 2007. ACM, 2007, pp. 155–162.

Worlds Apart - Doctors' Technological Frames and Online Medical Records

Åsa Cajander[1], Christiane Grünloh[2,3], Hanife Rexhepi[4]
[1] Uppsala University, Box 337, SE-751 05 Uppsala
[2] KTH Royal Institute of Technology, Lindstedsvägen 3, 10044 Stockholm
[3] CUAS Cologne University of Applied Sciences, Steinmüllerallee 1, 51643 Gummersbach
[4] University of Skövde, Box 408, 541 28 Skövde
asa.cajander@it.uu.se
christiane.gruenloh@fh-koeln.de
hanife.rexhepi@his.se

Abstract. The ability of individuals to access and use their online medical records serves as one of the cornerstones of national efforts to increase patient empowerment and improve health outcomes. However, the launch of online medical records in Uppsala County, Sweden, has been criticized by the medical profession and the local doctors' union. The aim of this paper is therefore to present the results from an exploratory study where interviews with two oncologists are analysed and discussed based on the theory of Technological Frames and Patient Empowerment. The results indicate that medical doctors have different assumptions and perspectives that affect their use of technology and how they view patient empowerment in everyday clinical work.

1 Introduction

The digitalization of health records enables the use of medical documentation in a different way than its original purpose. What started as a working tool for healthcare professionals to document and communicate patient-related information, can now also be used as a communication tool between the healthcare provider and the patient, since the latter can also access the medical record now. Other uses of the medical records are quality assurance, research and legitimizing the work done. Personalized patient information has potential benefits, e.g. increased sense of empowerment, improved patient satisfaction, improvements in patients' knowledge and understanding of their condition ([4], p. 83). Patient empowerment refers to a process with the purpose of educating people to be able to think critically and autonomously, where the outcome should be an enhanced sense of self-efficacy in the educated person [2]. In 2012 several eHealth services were deployed in Uppsala County, Sweden, including a service that enables citizens over 18 years of age to

access their medical records online. The deployment was preceded by several challenges [5] and especially the online medical records were criticized by the medical profession and the local doctors' union [6]. The medical professionals criticized that having patients accessing their medical records would endanger the effectiveness and value of the records, because these are viewed primarily as working tools ([5], [8]). Additionally the professionals viewed patients' access more as a means to control and monitor the professionals rather than a service to the patients [5]. Interestingly, patients already had the right to access their medical records before by requesting a paper copy sent by mail, but this has been considered "less interfering to the professionals' autonomy" [6]. However, not all professionals have been negative regarding patients' access, since some also considered it "valuable assistance as errors and mistakes are discovered and corrected" [6].

In this exploratory study, we analyse and discuss interviews with two oncologists from the theory of Technological Frames [7], which are composed of the assumptions and values that users have of technology. In the paper we have a special focus on the perspective of patient empowerment and the work environment of medical doctors, which to our knowledge has not been done before. The paper first presents the notion of Technological Frames and patient empowerment. This is followed by a description of the method. We then present the results from the analysis of the interviews from the perspective of Technological Frames and patient empowerment. This is followed by a discussion of the results in relation to other research and some implications for the introduction of eHealth systems. The purpose of the study is to spark a discussion and to receive feedback on these preliminary results.

2 Online Medical Records in Uppsala County Council

Introducing online medical records in Uppsala County Council in 2012 was the result of a process with different projects, law changes and pilot studies that started fifteen years earlier. One of the main reasons for launching online medical records in Uppsala County Council was to increase patient empowerment and to contribute to patientcentred care (for a full description of the development of the system in relation to laws and technical norms, see [5]). Medical records are available to patients through a national e-service called "My Healthcare Contacts". Patients access the service using an e-ID or alternative secure login options. When the system was launched in 2012, unsigned test results and medical records, i.e. information not yet approved by medical staff, were invisible to the patients. As soon as they were approved, patients could see them in their online record. This was the status of the

system when the data collection for this paper was made. However, in March 2013 this was changed and now patients can see test results and text in the online medical records as soon as they are entered into the system. However, the patient can see that these test results have not been verified by medical staff by a text in the system that says: *"This information is not signed, which means that it can be changed or deleted"*.

3 Theoretical Underpinnings

In this section we present the theory of Technological Frames [7], which is followed by a presentation of the concept of patient empowerment. The medical doctors' underlying assumptions and expectations are a part of their Technological Frame presented below. In the context of online medical records for patients, the underlying assumptions about patient empowerment are of special interest, since one of the goals when introducing the system was to improve patient empowerment.

3.1 Technological Frames

Orlikowski and Gash's theory on Technological Frames [7] is relevant to our study as it elucidates how different user groups interpret information technology (IT) differently and how these interpretations guide them to make sense and take action. These frames are social in nature, and have implications for technology development, implementation and use. Technological Frames constitute the "cognitions and values of users and designers" and "the underlying assumptions, expectations, and knowledge that people have about technology" [7]. In our study, the professionals working in healthcare have one interpretation of IT based on for example their purpose, context, knowledge base, power and previous experience of similar systems, which constitutes their technological frame. Orlikowski and Gash also propose the notion of congruence and incongruence in Technological Frames, where congruent frames are related to each other through their structure and content, whereas incongruent Technological Frames have important differences in expectations, assumptions or knowledge about some key aspects of technology which make these frames incompatible.

3.2 Patient Empowerment

Patient empowerment describes a situation where the patient's role is changing from a patronized patient to a patient that is informed, autonomous and

engaged in his or her own care [1]. The concept of empowerment is highly ambiguous and may be interpreted in many ways. Empowerment according to Shalom (2007) "involves a sense of control and self-efficacy, as well as an active position within the healthcare system" ([9], p.168), where one of the main sources for empowerment is information. Shalom also argues that there are different levels of patient empowerment [9]. On the basic level, the patient learns to ask questions relevant to his/her health condition. On an advanced level, the patient will become a partner in the decision-making process. Anderson and Funnell (2010) view empowerment as both a process and an outcome [2]. According to them, empowerment is a *process* when the purpose of an educational intervention is to increase one's ability to think critically and act autonomously. Empowerment is an *outcome* when an enhanced sense of self-efficacy occurs as a result of the process. They elaborate further that the empowerment approach involves that patients are facilitated and supported to reflect on their experience of living with a specific illness [2]. An important aspect of patient empowerment is also self-determination, which refers to the "philosophical view of humans having the right and ability to choose by and for themselves" [3]. At the same time, the empowerment-oriented approach also views patients as "being responsible for their choices and the consequences of their choices" [3]. This also includes determining which decisions they want to make themselves and when to ask healthcare professionals what to do [3]. While patients are in control of their daily self-management decisions and therefore responsible for the decisions they make, including the consequences, healthcare professionals are responsible for supporting patients to make informed self-management decisions [2]. However, Anderson and Funnell (2010) talk about healthcare professionals being socialized to a set of responsibilities and expectations, which might contradict the empowerment approach, but since it is deeply embedded in the professional identity, they might not even be aware of it [2]. As part of the empowerment process, not only the patients but also the professionals change, because they might have to "unlearn being in control" [3].

4 Method

This workshop paper presents the experiences of two oncology doctors when it comes to the launch of online medical records in Uppsala County Council, Sweden. These two oncologists were interviewed as part of a large interview study focusing on the medical professionals' perceptions and experiences of online medical records and other eHealth services. The interviews were conducted in the summer of 2013, which is about six months after the eHealth

services for patients were launched. For the purpose of this workshop paper, these interviews have been re-examined from the perspective of Technological Frames and patient empowerment. We chose to further examine the Oncologist interviews in this workshop paper, since, at the time of the launch of the eHealth services, Oncology was considered by many in the media as the most troublesome area of medicine from an ethical perspective. The two interviews were semi-structured and an interview template was used. The interviews were carried out on site and lasted for about one hour. One researcher conducted the interviews, and they were transcribed verbatim, but sometimes they have been slightly rephrased in this paper in order to be more readable. Moreover, we present the interviewees as women in this paper to make them anonymous. In the data analysis the following categories were used as a part of the thematic analysis [10], some of which were predefined and some emerged from the analysis of the data: What is their perspective and view of patient empowerment in connection to the launched eHealth services? What do they think about patients reading online medical records and test results? Have they changed their way of writing in the medical records? How does the eHealth Service online medical record affect the work environment? The interviews were read through and analysed separately by the three authors of the paper, and then we discussed and analysed them together. The writing of this workshop paper was also a part of the analysis (as presented by for example [11]).

5 Results

In this section we present findings related to the theme of Technological Frames and patient empowerment.

Patient Empowerment in Connection to the launched eHealth service

Oncologist A does not seem to consider patients' access to information (such as test results) as a way to increase patient empowerment. She believes that it is the doctor's responsibility to deal with test results. She emphasizes: *"They only get worried by reading, and they usually only focus on the medical details. They should live their lives as usual and come to the medical appointments without worrying. /.../ we try to do the best for all patients, /.../ it is we who take responsibility for complications and everything. Therefore they should try to enjoy life and not sit in front of the computer and check test results /.../. They're supposed to do other things, not to look for information online about which options they have."* When asked about the possible opportunities when patients are reading their medical records, Oncologist A

acknowledges that patients could prepare and read about alternative treatments. At the same time she is afraid that those well-read patients will be more demanding when they visit their doctor. Moreover, Oncologist A is worried about the future of her profession as a medical doctor when patients are able to access test results: *"Why do you need a doctor if patients themselves can look at the test results before they are signed /.../"* Oncologist B believes that empowering patients is important and she therefore encourages her patients to read their online medical records and other information on the Internet related to the patient's illness. Oncologist B also argues that patients today are more informed about their illness because of easy access to information. This has transformed the patient into an active collaborator in the patient meeting, where he/she asks questions relevant to his/her health condition, and where he/she takes part in the decision-making process: *"Based on the results we discuss what we should do or not do and together we discuss which treatments are appropriate /.../"* Oncologist B also believes reading their online medical record can help patients take responsibility of their own care and outcomes, which is important for achieving patient empowerment.

Perspective on Patients Reading Medical Records and Test Results

Oncologist A has an overall negative opinion about patients reading their online medical records: *"Well, I am negative to this system, even though it has its advantages, but at the end of the day I am negative."* Oncologist A has chosen not to sign the test results, because this will prevent patients from seeing this information. She feels upset knowing that patients will be able to access unsigned test results in the future without first consulting a doctor. She explains: *"Patients being able to see test results before consulting the doctor, it's a catastrophe!"* Oncologist B emphasizes that although she encourages patients to access their medical records, she also tells them that the access can have consequences, such as not being able to contact the physician immediately after reading the medical record. Oncologist B emphasizes: *"And then I tell them that they should access it, but that they also need to face the consequences of having to wait until their next planned patient visit to discuss the results. And I believe that this suits some patients, but many patients also say that 'I don't dare to read my medical record'. /.../ And I believe, even just read the notes about the disease contributes to improved engagements. So I am positive to this if it doesn't result in problems or harm for the patients, for example that they get notified at strange times when they cannot contact anyone. But when the patient is aware of this, then I don't see any problems with it. I mean you must take responsibility for your actions and if you want to log in and look for the test results at a certain time, then you have to take*

responsibility for it even if it's in the middle of the night." Oncologist B also believes that test results should be made available to the patient immediately without any delays (i.e. waiting for the doctors to sign them). Moreover, she considers the test results to be the most important information that patients want to have access to and therefore it should be made available immediately.

Changing the Way of Writing in the Medical Records

After the launch of online medical records, Oncologist A has become very careful about what to document in the medical records. She is especially careful when it comes to writing about progression, and she does not write anything in the medical records before she has informed the patient. Oncologist B argues that the online access to medical records has not had any major impact on documentation. However, she argues that she sometimes takes the opportunity to write a comment into the medical record to the patient, as she states: *"It is possible that I change a little bit my way of writing in the medical record, you can sneak in some messages to the patient such as quit smoking or something like that."*

Online Medical Records Affecting the Work Environment

Oncologist A has experienced that patients ask more detailed questions about their treatment, and specifically about different blood samples, after accessing their online medical records. She believes that access to the test results has a negative impact on the patient meeting as patients have a tendency to focus on the results of single samples, without understanding the full picture of the disease. She argues: *"It is just one test that the patient happens to see, and that test does not say anything about the health condition at large nor does it give a full picture"*. Oncologist A also believes that access to test results will contribute to increased phone calls with expectations to get quick answers regarding how the results affect the treatment. Furthermore, Oncologist A is worried about how patients accessing non-signed test results will affect her work environment. She claims that she will be more stressed knowing that the patients are reading the test results. She also feels that patients want to supervise and control doctors by reading their online medical record. According to her, this has negative implications on the work environment: *"We speak the truth so, it feels that the patients want to supervise us all the time, but we try to do our best, we do not work against patients"*. Oncologist B does not believe that the online medical record has had a major impact on the working environment. She argues that the concerns healthcare practitioners had regarding patients' access to test results seem to have been unfounded, as the latter has not been proven to have any adverse

effect. For example, the number of phone calls to the clinic has not increased. The physician emphasizes: *"Access to the online medical record can generate some phone calls but it could also take away some phone calls so therefore I do not think there will or has been an increased burden on the clinic, something that many of my colleagues have been worried about."*

6 Discussion and Conclusion

From the analysis of the interviews with Oncologist A and B we can see that they have incongruent Technological Frames, where their assumptions and views held about the function and role of the system are clearly different.

It seems that Oncologist A has a Technological Frame that focuses on the professional's perspective and their work environment. She does not mention patients in her reasoning. Oncologist B, however, has a clearer focus on the patient's perspective and patient empowerment. These two Technological Frames are incongruent when it comes to the assumption about the role of the system, where Oncologist A has the perspective that the role of the system is that of a healthcare communication tool. From Erlingsdottir (2014) we can learn that this perspective of Oncologist A seems to be the most common one at this point in time, as most doctors strongly feel that the role of medical records is to support the communication between healthcare providers [6]. Oncologist B has a Technological Frame where the system can also be used to communicate with patients and empower them.

It is noticeable that the Technological Frame connected to patient empowerment is different between the two doctors as well. Oncologist A does not seem to consider patient empowerment issues or that technology can facilitate patient engagement. She believes that the patients should focus on living their lives, trust their doctors and not read their medical records or other information online. Oncologist B sees the access to the online medical records as a way for patients to get information about risks related to their illness and as a way for them to take responsibility for their own care and outcomes, which is important for achieving patient empowerment. This concurs with Anderson and Funnell, who state that for example diabetes patients control and are responsible for 98% of their care, which conflicts with the socialization of healthcare professionals to take responsibility for the care and outcomes [2]. Oncologist B describes that she sneaks in comments in the medical records to affect the patient's decisions. This reflects a basic assumption that the medical doctor knows what is best for the patient, and it is not based on the view of a patient as an equal partner. This would be in line with the initial view of patient empowerment, i.e. that the underlying purpose

is to increase patient compliance or adherence [2]. However, according to Anderson and Funnell "Empowerment is the antithesis of compliance" and their intention with regard to patient education was to "increase the learner's freedom/autonomy (i.e. one's capacity to make informed decisions) rather than increase the learner's conformity / compliance (i.e., one's willingness to follow the instructions of those in authority)" [2]. It is noticeable from Oncologist A that online medical records are perceived as a threat to her professional role. In the interview it is clear that she is afraid that her job will become obsolete if patients have access to their test results. However, Anderson and Funnell (2010) emphasized the following with regard to diabetes patients: there is a need for two kinds of expertise (in their case: diabetes expertise by healthcare professionals and the equally important expertise of patients) and therefore collaboration is necessary [2].

Oncologist B is very aware of the stressful work environment of medical doctors, but she still thinks that it is a good idea that the patients read their medical records. She is aware that the work environment of the medical doctors would be very stressful if patients contacted the healthcare as soon as they have accessed their test result. Therefore, she expects patients to wait until their next visit to ask about specific test results. However, she does not think that the online medical record has had any negative impact on the work environment of doctors.

The results of our analysis support the view that the Technological Frame has implications for the use of a system. This can be seen when Oncologist A chose to *not* sign the results in order to prevent the patient from seeing them. In her view, patients should not bother reading their records. By choosing not to sign the test results, she puts her perspective on what is good for the patient into practice. This example also shows that how healthcare practitioners perceive technology affects how and whether they will try to make the patient an active and engaged collaborator. While Oncologist A uses the system to prevent her patients from accessing the test results, oncologist B is in favour of using technology as a way of increasing patient knowledge about their own health.

It is clear from this exploratory study on Technological Frames and patient empowerment that medical doctors have different assumptions and perspectives that affect their use of technology. These Technological Frames should be considered and discussed when implementing online medical records worldwide, especially when implementing eHealth services that aim to make the patient an active and central collaborator. Moreover, patient empowerment in relation to different eHealth services should also be discussed. If healthcare practitioners do not understand (1) how eHealth

services such as online medical records can be used to increase patient empowerment, and (2) how they themselves can support patient empowerment, the objective of the system will fail. Moreover, further research is needed on Technological Frames and their impact on the use and adoption of eHealth services.

Acknowledgements

The current research is part of the Swedish research projects Deployment of Online Medical records and E-health services (DOME) and the EPSA project. We are thankful for the time, support and access granted by SUSTAINS, co-funded by European Commission. The financial support of VINNOVA, the Swedish Governmental Agency for Innovation Systems, and AFA, are greatly acknowledged.

References

[1] Ammenwerth, E., Schnell-Inderst, P., Hoerbst, A.: Patient Empowerment by Electronic Health Records: First Results of a Systematic Review on the Benefit of Patient Portals. In: Studies in Health Technology and Informatics Volume 165. Amsterdam: IOS Press. pp. 63-67. (2011)

[2] Anderson, R. M., Funnell, M. M.: Patient Empowerment: Myths and Misconceptions. Patient Education and Counseling. 79(3), pp. 277-282. (2010)

[3] Aujoulat, I., d'Hoore, W., Deccache, A.: Patient empowerment in theory and practice: Polysemy or cacophony? Patient Education and Counseling. 66(1), pp. 13–20. (2007)

[4] Coulter, A.: Patient Engagement - What Works? Journal of Ambulatory Care Management. 35(2), pp. 80-89. (2012)

[5] Erlingsdottir, G., Lindholm, C.: When patient empowerment encounters professional autonomy: The conflict and negotiation process of inscribing an eHealth service. Scandinavian Journal of Public Administration (SJPA), pp. 1-33. (2014)

[6] Erlingsdottir, G., Lindholm, C., Ålander, T.: eHealth services, patient empowerment and professional accountability - An empirical study on the changing patient-doctor relationship in the digital world. Conference paper, Edinburgh. (2014)

[7] Orlikowski, W. J., Gash, D. C.: Technological Frames: Making Sense of Information Technology in Organizations. ACM Trans. Inf. Syst. 12(2), pp. 174-207, ACM, N.Y. (1994)

[8] Scandurra, I., Ålander, T.: Registered nurses' opinions of Patient Access to Electronic Health Records in Sweden - a Nationwide Survey. In: Medicine 2.0 '14 Congress. (2014)

[9] Shalom, N.: The health information specialist. A new role for channeling web information to promote patient empowerment. Journal of Information, Communication & Ethics in Society 5(2/3), pp. 167-184. (2007)
[10] Silverman, D.: Qualitative Research. 3rd. edition. Sage Publications Ltd. (2011)
[11] Wolcott, H. F.: Writing up qualitative research. Sage Publications, Inc. (2008)

Self-Help Obesity Prevention Program in Stokvels: A Social Media Intervention

W. Douglas Evans[1], Nelia P. Steyn[2], Marjanne Senekal[2]
[1]The George Washington University
[2]University of Cape Town
wdevans@gwu.edu

Abstract. Recent studies such as the South African National Health and Nutrition Examination Survey (SANHANES) indicate rates of overweight (26%) and obesity (41%) among the black South African female population, which represents a major public health threat. Social groups can be empowered as a channel for behaviour change. One group, the Stokvel, a private credit union in which black South Africans pool their money, offers a distinctive opportunity for capacity building and obesity prevention interventions. The Self-Help Obesity Prevention Program in Stokvels (SHOPPS) will offer a wellness program comprising a healthy diet, increased physical activity (PA), and a self-help cognitive-behavioral intervention for black South African women in Stokvels. Mobile phones and evidence-based, Web-based weight control programs are promising approaches for obesity prevention and SHOPPS will demonstrate their potential in conjunction with social support activities. This paper describes the SHOPPS conceptual framework, methods, protocol, and intervention strategy using digital media.

1 Introduction

The global obesity epidemic is an urgent public health problem [1]. Recent population-based obesity prevention efforts have applied an ecological model of health, which represents best practice [2]. Mobile health (mHealth) technologies are a new level of social ecology to affect obesity risk factors [3-4]. The current paper reports on a research protocol for a new study, the Self-Help Obesity Prevention Program in Stokvels (SHOPPS) intervention.

In September 2011, the United Nations (UN) resolution 65-238 called for new NCD prevention efforts [5]. South Africa's Strategic Plan for the Prevention and Control of NCDs 2013-17 specifically calls out NCDs as a threat to the country's health and health care in the next two decades [6]. Evidence of what interventions work and capacity building in local organizations and populations is needed to combat NCDs in South Africa.

The practice of unhealthy lifestyle behaviors such as inactivity, poor diet, stress, and overweight/obesity in black South African adult women, who represent a majority of *Stokvel* members nationwide, are detrimental to their

own health and communities, families and social groups. [7-8]. BMI values of black women indicate that 25 % are overweight (BMI = 25-29.9), 40.7% are obese (BMI>= 30), or 65.7% combined based on recent South African Health and Examination Survey (SANHANES) data [9]. This is a substantial increase from the 56.2% observed in a 2003 national survey [10]. Some 10.2% of black women were hypertensive; 24.9% had abnormal serum cholesterol; 45.4% had abnormal HDL-cholesterol values; 29.5% had abnormal LDL cholesterol values; 19.4% had abnormal triglycerides and 8.2% had HbA1c values above 6.5%. Obesity and sequalae (e.g., type 2 diabetes) have negative effects on adult South African's health [11-12].

Numerous studies have shown that health promotion in community and social group settings can improve adults' morale, health, fitness and dietary intake [13-14]. One reason for this may be that within-group social dynamics (e.g., support, mutual accountability, friendly competition) help group members to achieve personal or group health goals [15]. *Stokvels* have been defined by the president of the South African National Stokvel Association (SANSA), Andrew Lukhele as: "a type of credit union in which a group of people enter into an agreement to contribute a fixed amount of money to a common pool weekly, fortnightly or monthly" [16]. As social groups focused on finance and investments, the *Stokvels* offer social support for obesity reduction, supported by a technology intervention. Thus a wellness program for *Stokvel* members comprising a healthy diet, increased physical activity (PA), and a self-help cognitive-behavioral intervention may reduce or prevent obesity risk behaviors, and related chronic NCDs, as advocated by the World Health Organization (WHO) [17-18].

The *Self-Help Obesity Prevention Program in Stokvels (SHOPPS),* focuses squarely on the new NCD prevention priority in a country at ground zero for the global epidemic. *SHOPPS* will use the innovative approach of working with female-only *Stokvels* in the Johannesburg area through mobile and Web technology to promote healthy eating and active living (HEAL) and weight loss.

This paper describes the *SHOPPS* conceptual framework, methods, protocol, and intervention strategy using digital media. *SHOPPS* will use mobile and Internet technologies to provide cognitive-behavioral feedback to promote HEAL. The primary outcome of interest is weight loss of 5% among *SHOPPS* participants compared to control. We will conduct a cluster randomized controlled trial (RCT) of *SHOPPS*. Participants will get detailed, tailored feedback on nutrition and PA through their mobile phones and online regarding how to overcome barriers to HEAL behavior change in their community, and help to track progress in managing their weight. Tailoring

will occur through customized text messaging and online feedback based on individual progress. This will include making a personal resolution to lose 5% of their body weight, the target for this intervention, and will include a financial incentive for the *Stokvel* group (group reward if each member reaches her individual goal). Participants will use the Website to track personal progress and receive tailored feedback. We posit that increased engagement with the program, feedback, and creation of a program 'brand identity' to represent a 'trusted friend and advisor' will promote HEAL behaviors [19-20]. *SHOPPS* is based on Cognitive-Behavioral Therapy (CBT) [21-22], and Social Cognitive Theory (SCT) [23]. Outcomes include changes in HEAL behaviors and weight loss.

Numerous studies have shown that health promotion in community and social group settings can improve adults' morale, health, fitness and dietary intake [24-27]. One reason for this may be that within-group social dynamics (e.g., support, mutual accountability, friendly competition) help group members to achieve personal or group health goals [28]. Thus a wellness program for *Stokvel* members comprising a healthy diet, increased physical activity (PA), and a self-help cognitive-behavioral intervention may reduce or prevent obesity risk behaviors, and related chronic NCDs, as advocated by the World Health Organization (WHO) [29-30]. There are no similar published interventions in developing countries [31].

Mobile phones are poised to be a powerful tool to promote health in South Africa. As reported by the United Nations Foundation (2009), Project Masiluleke in South Africa reaches up to one million people each day with HIV prevention, testing and treatment information [32]. Some mHealth studies on weight loss, nutrition and PA promotion in developed countries have shown promise. Patrick and colleagues (2009) found, that compared to control, SMS and Multimedia Message Service (MMS) participants adjusted average weight loss was 1.97 kg greater [33]. Hurling and colleagues (2007) found that an Internet and mHealth intervention among overweight adults that included reminders produced over 2 hours more PA per week compared to adults with no access [34]. Joo and colleagues (2007) found that weekly text messages about diet and PA behavior promoted weight loss [35]. Our use of SMS in *SHOPPS* is significant in that it will extend the evidence base in mHealth and obesity prevention to developing countries [36].

2 Methods

We will evaluate *SHOPPS* in a cluster RCT. A total of 22 *Stokvels* (average 20 members per *Stokvel*) will be randomized to a *SHOPPS* treatment group

and a control group (11 *Stokvels* in each). *Stokvel* members in both conditions will receive *SHOPPS* plus standard printed South African Department of Health (DOH) information on nutrition, PA and wellness. Participants assigned to the control group will receive the DOH information only. We will recruit *Stokvels* based on: 1) comparable size (15-25 members), 2) located in the Johannesburg metro area, and 3) all black African female members.

Additionally, the protocol requires that all individual participants have BMI of >25 kg/m^2 (i.e., overweight or obese). The goal is to maximize internal validity rather than external validity. As the *Stokvel* is the sampling unit, all members of a group will be in one study condition. We will include all members of a *Stokvel* that meets group criteria *in the intervention activities*. However, investigators will separately analyze those members who meet minimum individual BMI criteria for outcome evaluation purposes. This will enable full group participation by all members, but maintain intent-to-treat principles for analysis.

Aim 1: Conduct formative research with African women to tailor *SHOPPS* to the target population. Hypothesis (H) 1.1: Participants will have positive message receptivity to *SHOPPS* on validated scale [37]. Aim 2: Conduct a 6-month cluster RCT of *SHOPPS* in Johannesburg area *Stokvels*. H2.1: Compared to control, *SHOPPS* participants lose 5% more weight after 6 months. H2.2: *SHOPPS* brand equity will mediate intervention effects on HEAL and weight loss outcomes.

There are three sources of data for the cluster RCT: 1) Anthropometric data collected through onsite physical examinations; 2) self-report through an in-person questionnaire at baseline and after 6 months and 3) online recording of weight status and participation in intervention activities (eg, providing shopping lists). Note that the research team will test for serostatus, and HIV-positive women will be eligible. We anticipate that any HIV-positive women will have already begun HIV treatment before any possible wasting (weight loss) and *Stokvel* participation.

To ensure comparability and sufficient numbers of participants, the research team will only sample from Johannesburg *Stokvels*. The research team has a list of Johannesburg registered *Stokvels* from the South African National Stokvels Assocation (SANSA), which provides background information on these groups. Only *Stokvels* meeting inclusion criteria will be included. This will ensure relative homogeneity of *Stokvel* group environments.

The research team has calculated the required sample size for analysis of covariance in a cluster-randomized trial [38]. We set the two-sided significance level at 0.05 and power at 0.80. We assume a baseline mean (SD) of weight of 75 (18.0) kg. These are based upon figures from the recent

SAHANES [9], with minor adjustments reflecting the exclusion from our analyis of women with BMI<25 kg/m^2. We assume an intraclass correlation coefficient of 0.02 and refusal and attrition of 10%, and we estimate that the typical number of eligible women per *Stokvel* will be 13 (based on the 65.7% national average overweight/obesity among black African women and 20 members per *Stokvel*). Based upon temporal stability estimates [39-40], the six-month stability of weight will be 0.85. Finally, following previous studies, we seek to detect a 5.0% weight loss among treatment participants [41]. Thus the required baseline sample size is 131 participants, or 11 *Stokvels* averaging 13 members. This is similar to other web-based weight-loss studies [42].

Before the recruitment, the research team will visit each sampled *Stokvel* at their monthly meeting in person in the Johannesburg metro area and ensures that all group members fully agree to terms of the study and signs a written agreement to participate (part of the IRB protocol). Next the research team will conduct a second visit to the *Stokvel* at their subsequent monthly meeting, conduct group orientation of all members, describe the study using a pre-developed script, and recruit members. An initial 24-hour dietary intake recall will be done at this time, on a weekday. This will be conducted as a fun event with refreshments. At the second visit, the baseline assessment, including all instruments and a second 24-hour dietary intake recall, will be done, all on a weekend day. After the baseline interview, all participants will receive the SA DOH basic dietary guidelines printed brochure, a $5 equivalent gift card, and a t-shirt. Then the intervention will begin and *SHOPPS* participants will receive password-protected access to the Website. After 6 months, the team will follow up using same procedures and participants will receive another $5 gift card.

After six months, the research team will return to each *Stokvel* and administer the follow up assessments. We will administer a physical exam, as in the baseline assessment. The interviews will include a supplemental module of questions on intervention exposure and reactions with the two treatment groups. The supplemental module will be used to measure self-reported dosage of *SHOPPS* received and satisfaction/preferences for specific intervention components. All interviews will be conducted in person by the investigators. To ensure retention, we will contact a representative of the *Stokvel* by phone, e-mail, and text one week before follow-up. Participants will be tracked via unique case ID for confidentiality.

3 Discussion

SHOPPS will be tailored both to the individual's HEAL and weight loss goals and to their success or specific needs as they progress. Tailoring will be based on a) individual response to keywords through the mProve texting platform (based on an algorithm programmed by mProve that triggers individually tailored feedback texts) and b) information provided by participants , and goal progress captured through the *SHOPPS* Website.

The Website includes detailed information, self-help advice, and progress monitoring tools on diet, PA, and weight status. The program includes a Cognitive Behavioral Therapy (CBT) component to assist the participant in adopting positive behaviors in order to follow the program successfully. The Website comprises the following sections of information: 1) Health benefits of weight loss; 2) explanation of a reasonable body weight and what it comprises; 3) methods of changing current eating and PA behaviors; 4) a healthy eating plan which can be individualized; 5) a PA guide; 6) weight loss strategies; 7) maintaining weight loss and a PA program; 8) weight, diet and PA self-monitoring tools. Participants will be required to monitor their weight status and record this on a weekly basis. The Website also includes weekly activities which the participant follows in order to engage actively with every aspect of the program. The website will also link to numerous other wellness Websites, including those dealing with stress and tension. As the participants follow the program from week 1 through to week 26, they will receive feedback on physical activity and weight status. We hypothesize they will become engaged in *SHOPPS* and build brand equity, which mediates weight status and other outcomes.

Texts will be delivered through a platform that supports interactive SMS using a content management system that delivers texts. The *SHOPPS* protocol will be an application on the platform that captures keyword responses to texts received and includes an algorithm that calculates a tailored *SHOPPS* text response based on the participants' history. The research team will design a branching tree structure of tailored texts to be delivered based on participant response. For example, if a participant has a history of replying to keywords for physical activity texts, they may receive in-depth suggestions for convenient daily exercise options in the local community on weekdays and weekends. Thus the SHOPPS protocol will provide tailored text feedback and complement tailored website feedback on progress.

We will use a "group contest" approach with a prize equal to the annual contribution for a member of the *Stokvel* if all overweight (BMI>25) group members reach the outcome of 5% weight loss. Based on information

provided by SANSA, an average *Stokvel* member in Johannesburg contributes approximately $200USD equivalent (R2400 in South African Rand) to the group each year. If each overweight member of the group meets the goal of 5% weight loss, the research team will deposit $200USD into the group's collective *Stokvel* account, proportionally increasing the value of each member's share of the account. The contribution will only be made if every overweight group member meets her goal, providing added incentive for achievement of intervention goals according to behavioral economic theory [43-44]. All members will benefit from the incentive, including those who were not overweight at baseline, as it is a group benefit.

4 Conclusion

The study will build capacity in mobile health in South Africa with University and community partners, extend the evidence on use of digital technology for weight loss, and reach women at high risk of obesity and NCDs. It is highly scalable, and if shown to be efficacious, will provide a basis for social group based obesity interventions using digital media with other high risk populations in South Africa and other low and middle income countries.

References

[1] World Health Organization. Global strategy on diet, physical activity and health: a framework to monitor and evaluate implementation. Accessed 23 June 2015 at: http://www.who.int/dietphysicalactivity/M&E-ENG-09.pdf. 2008.

[2] Evans, W.D., & Hersey, J. Community programs to prevent obesity. In *Encyclopedia of Obesity*, Golson, J.G. and Keller, K. (Eds.), Thousand Oaks, CA: Sage Publications, Inc., 2010.

[3] Abroms, L.C., Padmanabhan, N., & Evans, W.D. (2012). Mobile Phones for Health Promotion. In *Interactive Health Communication Applications: Promising Strategies for Health Behavior Change,* S. Noar, and N.G. Harrington (Eds.). New York: Routledge, 2012.

[4] Evans, W.D., Abroms, L.C., Poropatich R., Nielsen, P., Wallace, J. Mobile health evaluation methods: Opportunities, Challenges and the *text4baby* Case Study. *Journal of Health Communication,* 2012, 17:sup1, 22-29.

[5] Beaglehole, R., Bonita, R., Horton, R., et al., for the *The Lancet* NCD Action Group and the NCD Alliance. Priority actions for the non-communicable disease crisis. *The Lancet*, 2012, 377(9775):1438-1447, doi:10.1016/S0140-6736(11)60393-0.

[6] South African National Department of Health. Strategic Plan for the Prevention and Control of Non-Communicable Diseases 2013-17. Accessed 24 June 2015 at: http://www.health-e.org.za/2013/09/19/strategic-plan-prevention-control-non-communicable-diseases-2013-17/. 2012.

[7] Steyn, N.P., Nel, J.H., Dhansay, M.A., Tichelaar, H.Y., Kriek, J.A., Prinsloo, J.F., & Benadé, A.J.S. Differences in nutritional status between caretakers of underweight and normal weight Pedi pre-school children. *South African Journal of Food Science and Nutrition,* 1995, 7: 53-59.

[8] Steyn, N.P., Lambert, E.V., Parke, W., Mchiza, Z., & De Villiers, A. A review of school nutrition interventions globally as an evidence base for the development of the HealthKick programme in the Western Cape, South Africa. *South African Journal of Clinical Nutrition,* 2009, 22(3):145-152.

[9] Shisana, O., Simbayi, L., Labadarios, D., et al. The South African Health and Examination Survey, 2012. Pretoria: HSRC. Accessed 22 June 2015 at: http://www.hsrc.ac.za/uploads/pageContent/1906/Vol%206%20No%201%20May%202012.pdf

[10] Department of Health, Republic of South Africa. Demographic and Health Survey, 2003, Full Report. Accessed 22 June 2015 at: http://www.measuredhs.com/pubs/pdf/FR206/FR206.pdf.

[11] Pelletier, K.R. Clinical and cost outcomes of multifactorial, cardiovascular risk management interventions in worksites: a comprehensive review and analysis. *Journal of Occupational and Environmental Medicine,* 1997, 39(12):1154-69.

[12] Pelletier, K.R. A review and analysis of the clinical and cost effectiveness studies of comprehensive health promotion and disease management programs at the worksite:1998-2000 update. *American Journal of Health Promotion,* 2001, 16(2): 107-16.

[13] Anderson, A.S., Cox, D.N., McKellar, S., Reynolds, J., Lean, M.E., Mela, D.J. Take Five, a nutrition education intervention to increase fruit and vegetable intakes: impact on attitudes towards dietary change. *British Journal of Nutrition*, 1998, 80(2):133-40.

[14] Aldana, S.G., Greenlaw, R.L., Diehl, H.A., Salberg, A., Merrill, R.M., Ohmine, S., Thomas, C. Effects of an intensive diet and physical activity modification program on the health risks of adults. *Journal of the American Dietetic Association*, 2005, 105(3):371-81.

[15] Marteau, T.M., Ashcroft, R.E. Using Financial Incentives to Achieve Healthy Behaviour. *British Medical Journal,* 2009, *338:b1415*.

[16] African Response (2012). Stokvels- A hidden economy. Accessed 23 June 2015 at: http://www.africanresponse.co.za/PressReleases/Documents/2012StokvelHiddenEconomy.pdf

[17] World Health Organization. WHO global strategy on diet, physical activity and health: a framework to monitor and evaluate implementation. Geneva, WHO. 2008.
[18] World Health Organization. Preventing chronic diseases-a vital investment. Geneva: WHO. 2005.
[19] Puoane, T., Tsolekile, L., & Steyn, N.P. Perceptions about body image and sizes in black African girls living in Cape Town. *Ethnicity and Disease*, 2010; 20: 29-34. http://www.ishib.org/journal/20-1/ethn-20-01-29.pdf
[20] Evans, W.D., Christoffel, K.K., Necheles, J., Becker, A.B., & Snider, J. Outcomes of the *5-4-3-2-1 Go!* Childhood Obesity Community Trial. *American Journal of Health Behavior,* 2011, 35(2):189-198.
[21] Shaw, K., O'Rourke, P., Del Mar, C., & Kenardy, J. Psychological interventions for overweight or obesity. Cochrane Database Systematic Reviews, 2005, Apr 18;(2):CD003818
[22] Cooper, Z., Doll, H.A., Hawker, D.M., Byrne, S., Bonner, G., Eeley, E., O'Connor, M.E., & Fairburn, C.G. Testing a New Cognitive Behavioural Treatment for Obesity: A Randomized Controlled Trial with Three-Year Follow-up. *Behaviour Research and Therapy,* 2010, 48(8):706-713.
[23] Bandura, A. Health promotion by social cognitive means. *Health Education & Behavior,* 2004, **31**(2):143–164.
[24] Simkin-Silverman, L.R., Wing, R.R., Boraz, M.A., Kuller, L.H. Lifestyle intervention can prevent weight gain during menopause: results from a 5-year randomized clinical trial. *Annals of Behavioral Medicine,* 2003, 26(3):212-20.
[25] Anderson, A.S., Cox, D.N., McKellar, S., Reynolds, J., Lean, M.E., Mela, D.J. Take Five, a nutrition education intervention to increase fruit and vegetable intakes: impact on attitudes towards dietary change. *British Journal of Nutrition*, 1998, 80(2):133-40.
[26] Aldana, S.G., Greenlaw, R.L., Diehl, H.A., Salberg, A., Merrill, R.M., Ohmine, S., Thomas, C. Effects of an intensive diet and physical activity modification program on the health risks of adults. *Journal of the American Dietetic Association*, 2005, 105(3):371-81.
[27] Burke, V., Giangiulio, N., Gillam, H.F., Beilin, L.J., Houghton, S. Changes in cognitive measures in a randomized controlled trial of a health promotion program for couples targeting diet and physical activity. *American Journal of Health Promotion,* 2004, 18(4):300-11.
[28] Marteau, T.M., Ashcroft, R.E. Using Financial Incentives to Achieve Healthy Behaviour. *British Medical Journal,* 2009, *338:b1415*.
[29] World Health Organization. WHO global strategy on diet, physical activity and health: a framework to monitor and evaluate implementation. Geneva, WHO. 2008.
[30] World Health Organization. Preventing chronic diseases-a vital investment. Geneva: WHO. 2005.

[31] Vital Wave Consulting. mHealth for Development: The Opportunity of Mobile Technology for Healthcare in the Developing World. Vodaphone and the United Nations Foundation. 2009.

[32] Patrick, K., Raab, F., Adams, M., Dillon, L., Zabinski, M., Rock, C.L., Griswold, W., & Norman, G.J. A text message-based intervention for weight loss: Randomized controlled trial. *Journal of Medical Internet Research,* 2009, 11(1) 1-9.

[33] Hurling, R., Catt, M., Boni, M. D., Fairley, B. W., Hurst, T., Murray, P., et al. Using internet and mobile phone technology to deliver an automated physical activity program: randomized controlled trial. *Journal of Medical Internet Research,* 2007, 9(2), e7.

[34] Joo, N.S., & Kim, B.T. Mobile phone short message service messaging for behaviour modification in a community-based weight control programme in Korea. *Journal of Telemedicine and Telecare,* 2007, 13(8):416-20.

[35] Riley, W.T., Rivera, D.E., Atienza, A.A., Nilsen, W., Allison, S.M., & Mermelstein, R. Health behavior models in the age of mobile interventions: are our theories up to the task? *Translational Behavioral Medicine,* 2011, 1:53–71, doi: 10.1007/s13142-011-0021-7

[36] Institute of Medicine. *Weight the Options: Criteria for Evaluating Weight Management Programs. Washington, DC: National Academy Press. 1995.*

[37] Petty, R.E., & Cacioppo, J.T. The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology,* 1986, 19, 123–205.

[38] Teerenstra, S., Eldridge, S., Graff, M., de Hoop, E., & Borm, G.F. A simple sample size formula for analysis of covariance in cluster randomized trials. *Statistics in Medicine*, 2012, 31:2169-2178, doi: 10.1002/sim.5352.

[39] Herman, K.M., Craig, C.L., Gauvin, L., & Katzmarzyk, P.T. Tracking of obesity and physical activity from childhood to adulthood: The Physical Activity Longitudinal Study. *International Journal of Pediatric Obesity*, 2009, 4:281-288, doi: 10.3019/17477160802596171.

[40] Katzmarzy, P.T., Pérusse, L., Malina, R.M., Bouchard, C. Seven-year stability of indicators of obesity and adipose tissue distribution in the Canadian population. *American Journal of Clinical Nutrition*, 1999, 69:1123-1129.

[41] Christian, J.G., Tsai, A. G., & Bessesen, D.H. Interpreting weight losses from lifestyle modification trials: using categorical data. *International Journal of Obesity,* 2010, 34**,** 207–209, doi:10.1038/ijo.2009.213.

[42] Neve, M., Morgan, P.J., Jones, P.R., & Collins, C.E. Effectiveness of web-based interventions in achieving weight loss and weight loss maintenance in overweight and obese adults: a systematic review with meta-analysis. *Obesity Review*, 2010, 11(4):306-21.

[43] Kahneman, D., & Riepe, M.W. Aspects of Investor Psychology, *Journal of Portfolio Management,* 199998, 24:52-65.

[44] Barberis, N. & Thaler, R.H. A survey of behavioral finance, In *Handbook of the Economics of Finance*, Constantinides, G., Harris, M. & Stulz, R. (Eds.), North-Holland, Amsterdam. 2003.

MyData Collection for Personal Health: Concept Design of a Lifestyle App for Junior Athletes

Jonna Häkkilä[1], Mira Alhonsuo[1], Juho Rantakari[2], Ashley Colley[2], Lasse Virtanen[2]

[1]Faculty of Art and Design
University of Lapland
96400 Rovaniemi, Finland

[2]Center for Internet Excellence
University of Oulu
90014 Oulu, Finland

[1]Firstname.lastname@ulapland.fi
[2]Firstname.lastname@cie.fi

Abstract. Collecting a digital footprint of data from our everyday activities is becoming an information source for preventive health care. Wearable sensor technologies combined with mobile phone applications offer an interesting way to collect and monitor personal activity data for personal use, as well as for wellness and health care professionals. In this position paper we argue for this approach, and present a concept design for a mobile app targeted at junior athletes, which aims to combine different possibilities of mobile technologies to create a tool that can provide versatile support for wellness.

1 Introduction

Tomorrow's health related systems and services will increasingly take advantage of a myriad of different sensor systems, which can track our physical activity and everyday life. Today, sensor technologies have achieved technical feasibility, miniaturization level and cost efficiency such that they can be easily integrated to various types of everyday objects. This enables omnipresent tracking of our activities, which consequently provides an overview of our lifestyle. Approaches such as described by the MyData model [5], highlight the importance of personal ownership of health data, and become critical as the number of sources and potential output channels for an individual's health data increase.

Wellness gadgets that collect data on our sports performance, such as heart rate monitors and step counters, have been available for large audiences for many years. Sensors integrated to clothes and other wearable form factors are about to take off in a large scale and result in increases in the amount of data collected. Today, these trackers and other wellness gadgets are typically combined with a PC or smart phone application where the user can view

his/her performance. Especially, the smart phone is an interesting platform because of its form factor, mobility and connectivity features. Ahtinen et al. [1] report that one of the key benefits the users saw in a mobile phone wellness tracking app was the fact that, as they carried the phone with them anyway, integrating the tracking technology to the routines was easy.

However, when investigating 39 popular (most downloaded) mobile wellness apps, it was found that most applications were still lacking guidance and target setting features [4]. There is still much potential that can be developed further in the features as well as in the user experience (UX) design of mobile wellness technologies. In [3], it is concluded that the key design requirements for technologies that encourage physical activity are 1) Give users proper credit for activities, 2) Provide personal awareness of activity level, 3) Support social influence, and 4) Consider the practical constraints of users' lifestyles.

In this positioning paper, we present our concept designs for mobile apps targeted to the persona of an active youth, more specifically an ice-hockey playing teenager. With this concept, we aim to demonstrate a holistic view of collecting and presenting wellness data that forms part of the MyData ecosystem and contributes to the user's overall health and wellbeing.

2 Mobile Health Application Concepts for Active Youth

Our design process began with a workshop following Service Design methodologies. Here, the preliminary aim was to understand the different stakeholders related to our defined persona, a 13 year old athlete / ice hockey player we named 'Niklas'. The output identified a large number of stakeholders, who could potentially provide and support the young person's everyday wellbeing. In a follow-up workshop including researchers and industry participants (Figure 1), a day-in-the-life of our young athlete persona was constructed. The output materials from the Service Design workshops enabled us to create three concept designs for mobile apps for young athletes: MyData Training app, MyData Healthy Eating Guidance app and Ask! app.

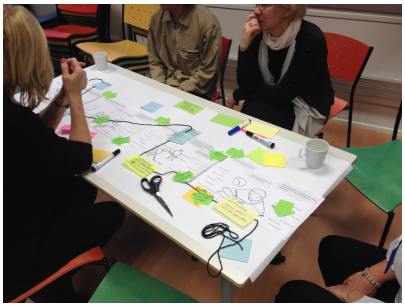


Figure 1. Creating the service concept in the service design workshop.

Prototypes of the three application concepts were made and a video showcasing the concepts in their envisioned usage scenarios was created, see Figure 2. The concepts videos introduce scenarios from the daily life of our persona 'Niklas', and show how various aspects of personal data collection can be used to enhance his training experience and improve his physical and mental health. These applications drew together the input from workshops, and were designed to fit to different aspects of the persona's daily life. In our ongoing research, the concepts will be used as experience probes.

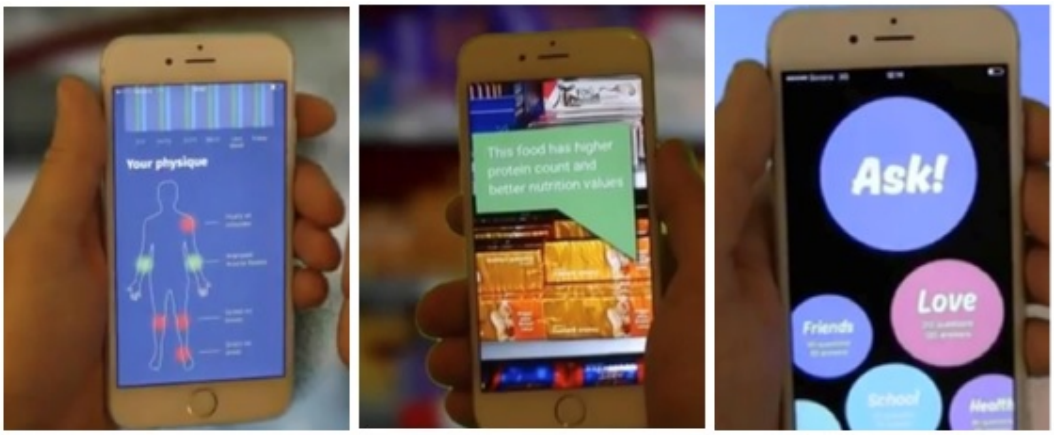


Figure 2. Left: The MyData Training app concept. Center: The MyData Healthy Eating Guidance mobile app concept. Right: The Ask! App concept [6].

The first concept is a Training app where Niklas can check his performance after the training session and compare it to another players and teams (Figure 2, left). Activity data is collected from sensors built into the player's hockey gear. The information is accessible to the user themselves, but also to their coaches, health professionals and parents if sharing has been agreed, following MyData guidelines.

The second application concept is a Healthy Eating Guidance app, where Niklas gets personalized nutrition advice when shopping (Figure 2, center). The App identifies products on the supermarket shelves and gives the user

recommendations what to eat. Niklas can also take photos of his meals and compare their nutritional value to that of his sports idol's meals. Hence we aim to guide the user towards healthier eating habits.

The third concept was created to answer user's mental issues and questions rising from puberty. The Ask! concept (Figure 2, right) gives guidance on mental wellbeing related issues for example stress, school or friends. Guidance is provided by counselors who are experts in their area. The user can also browse previously asked questions and answers. The Ask! concept gives users easy way to get personal help from professionals without the feeling of embarrassment.

3 Discussion

In this paper, we have presented our on-going work, where we position ourselves to the approach that the creation and management of an individual's MyData digital health footprint can be used to create holistic health and wellness services. By tracking everyday life activities, we can provide services that are integrated to our everyday lives, and create individually focused awareness and information to enable individuals to pursue a healthier lifestyle. It is to be noted that our research does not aim to new mobile application concepts, but we target to use the concepts as probes in our further user research. Here, we address an unexplored user group, the active youth. Whereas the quantified-self user group [2] is interested in following their wellness data closely and in great detail, this may not generalize to other groups. In tracking applications, several challenges rise e.g. from the basis of data sharing and privacy, as well as detection accuracy and reliability of the interpreted results. Here, one interesting aspect of our selected user group is that their view of data privacy may differ from that of a more typically studied group.

Acknowledgements

This research has been supported by a grant from Tekes – the Finnish Funding Agency for Innovation as part of Digital Health Revolution programme.

References

[1] Ahtinen, A. Isomursu, M. Huhtala, Y., Kaasinen, J., Salminen, J., Häkkilä, J.: Tracking Outdoor Sports – User Experience Perspective. In Proc. AmI'08. Springer-Verlag (2008).

[2] Choe, E. K., Lee, N. B., Lee, B., Pratt, W., Kientz, J. A. Understanding quantified-selfers' practices in collecting and exploring personal data. In Proc. CHI'14, 1143-1152. ACM (2014).
[3] Consolvo, S. Everitt, K., Smith, I., Landay, J.A.: Design Requirements for Technologies that Encourage Physical Activity. In Proc. CHI'06. ACM (2006).
[4] Häkkilä, J., Colley, A., Inget, V., Alhonsuo, M., Rantakari, J. (2015). Exploring Digital Service Concepts for Healthy Lifestyles. In Proc. HCI International 2015, Springer.
[5] http://www.lvm.fi/julkaisu/4440204/mydata-a-nordic-model-for-human-centered-personal-data-management-and-processing (Accessed June 1, 2015)
[6] https://www.youtube.com/watch?v=GPmClDffpTo (Accessed June 11, 2015)

The Future of Digitally Enabled Health Coaching – A Proposed Model

Charalampos Kyfonidis, Marilyn McGee-Lennon
Computer and Information Sciences
University of Strathclyde
26 RICHMOND STREET G1 1XH, Glasgow, UK
[charalampos.kyfonidis;marilyn.lennon]@strath.ac.uk

Abstract. The epidemic of chronic diseases has started worrying health bodies. The costs of dealing with such a problem (time, money and personnel) are continuously increasing especially with an ageing population. This paper proposes a conceptual large scale digital health coaching intervention model that could be applied by any public health body to design, implement or rationalise digital health coaching solutions. The model aims to support the increase of patient empowerment and the decrease of costs by redistributing the available resources more efficiently across the whole eco-system with the use of Artificial Intelligence coaches. The model is not targeting the replacement of human presence by computers, but a coaching strategy that will enable, assist, promote interaction and help automate (where needed and possible) resource consuming processes.

1 Introduction

In 2006 the World Health Organisation referred to chronic diseases as a "global epidemic"[1]. In 2012 chronic diseases were the leading cause of mortality with more than 21 million deaths worldwide[2]. In the UK, people with long term conditions account for 50% of all GP appointments, 70% of all inpatient bed days and 70% of overall NHS's expenses [29, 37]. As a consequence, health-care bodies have to spend immense resources (time, money and personnel) in order to ameliorate the impact of treating patients with long term conditions and diseases. This epidemic of chronic diseases has a direct impact on the economy and increasingly is a plausible concern from governments, public bodies and global organisations [41] on how these rising costs could be mitigated.

According to Golubic [16], the main reason behind the development and progression of preventable chronic diseases is unhealthy lifestyle. People with diseases such as arthritis, diabetes, obesity, cardiovascular disease,

---

1 WHO 2006 - Chronic disease handbook: http://www.who.int/chp/advocacy/en/

2 WHO 2014 - The top 10 causes of death : http://www.who.int/mediacentre/factsheets/fs310/en/

osteoporosis, arrhythmia etc. should have a healthy diet and exercise on a regular basis. Solutions therefore must focus on encouraging and supporting health behaviours and lifestyle choices. The unsustainability of modern -western- healthcare systems is also partly a result of the tendency to focus in the treatment of diseases rather than their prevention [18]. Since a patients' condition is dependent largely on their self-care and well-being [13], motivation is an essential aspect of their treatment plan and solutions need to address this. In 95% to 99% of those cases the health conditions are managed by the patients themselves [22] and unfortunately in most cases, patients lack a full understanding of their condition and how to manage it [4].

The goal of this position paper is to provide some insights on how the well-being of chronic disease patients can be enhanced by a more integrated and holistic health coaching model. This work tries to eschew the reactionary approach of cutting costs without caring about the whole ecosystem; instead it proposes a conceptual intervention model that can be used to balance the goals of treating patients as individuals and promoting the sustainability of health-care systems.

2 Background

2.1 Persuasion and Behaviour Change

Behaviour change is the ultimate goal of many health and wellness interventions and programmes. It is achieved through continuous feedback cycles where the subject is setting goals, identifying competitive goals and resetting goals by taking into account current status [8]. Behaviour change is hard to achieve, even if you manage to change someone's attitude [17] and can involve methods and techniques to persuade or motivate the user to choose a particular behaviour..

Persuasion is a way to provide incentives. It is a social interaction that consists of two social entities and a stimuli-message [21]. The main idea behind persuasion is to provide motivation and ideally influence the subject to abandon one set of behaviours and to adopt another [27]. As stated in the literature [14], persuasion can be achieved through social cues and strategies like reduction, tunnelling, tailoring, suggestion, self-monitoring, surveillance and conditioning.

There are three different types of persuasion [19]:

1. Interpersonal persuasion: When someone tries to persuade another individual through personal interaction.

2. Computer-Mediated persuasion: When someone uses a computer in order to persuade another individual (e.g. online advertisements).
3. Human-Computer persuasion: When an artificial-computer agent tries to persuade a human (e.g. health coaching apps such as UbiFit [10]).

Persuasion is a theory that has been vastly researched and a lot of time and effort is spent towards its use as a countermeasure for the impacts of chronic diseases. Persuasion is the principal behind almost every health coaching intervention.

2.2 Health Coaching

A common practice for helping patients to achieve their health-related goals by enhancing their well-being is called health-coaching [32]. Even though there is no unanimity on the definition of what health coaching is [40], a good definition is " *a practice framework that complements patient teaching and supportive therapy as a method for enhancing self-care and self-management behaviour for people with chronic disease and their family members*" [25].

Health-coaching is based on the principle of authority [14], according to which people tend to defer to authorities [9]. People, by presuming that authorities are knowledgeable and powerful, expect their guidance, recommendations and helpful information [14]. It is a purely patient-centred approach and it is based on the interpersonal relationship of the patient with the coach. The coach in that context is a professional educated and experienced in behaviour change in health matters [39].

By taking advantage the aforementioned facts, a lot of health coaching programmes have been established and running in US and UK and have instituted it as an important part of chronic conditions' management [38]. In those programmes clinicians are educating, encouraging and helping patients to acquire skills and tools in order to actively participate in their care. Hence, they can manage their condition and reach their self-defined health goals [3].

Unfortunately, even though this approach has multiple benefits, for both clinicians and patients [30], and is very effective [3] it is also very expensive [38]. Under normal circumstances, primary care clinicians spend 15 minutes for every patient and are striving to fit multiple agenda items into that time [31]. Hence, they cannot meet the needs of the numerous chronic conditions patients and thus cannot effectively coach them. Moreover, it is very difficult to scale up this kind of coaching [38] because a lot of time and resources are needed in order to recruit and train clinicians.

2.2.1 Remote Health Coaching

Remote coaching is defined as any coaching interaction that takes place from distance [33]. Remote coaching models are proven to be beneficial and can replace regular face-to-face coaching [34]. Mobile phones can facilitate remote coaching interventions and patients can easily collect data about their activities and physiological measures [11] allowing the feedback needed to be concrete and targeted to their particular case. Nonetheless, the increasing number of patients [30], demands further automation of health coaching. Consequently, many researchers have proposed and implemented computer agents as coaches.

2.2.2 Artificial Agent Coaching

A health coach does not necessarily need to be human. One of the main advantages of artificial agents- compared to humans is their ability to automate processes. The automation of the decision support process, when it comes to integrated health-care, can be a very helpful tool [18] and reduces costs [24]. It can be away of providing, personalised information for the patient (by tailoring evidence from the literature in his profile), reasoning support, guidelines and instructions [18]. Automation can also however introduce new difficult problems such as (i) issues of (real or perceived) responsibility and risk when replacing human interactions with automation, (ii) issues of (real and perceived) privacy and security when patient information flows are changed and (iii) issues of feasibility of implementing such automated frameworks in terms of what (artificial intelligence) technologies already exist and what still needs further research and development in order to fully support self-management.

2.3 Empowerment and Peer Support

The main goal of every health coaching system, independently of the intervention type, should be the empowerment of the patients. The term patient empowerment describes the augmented ability of patients to proactive perceive, impact and control their own health status [7]. Another interesting approach for the management of chronic conditions is peer support. Peer support occurs when people who have the same condition provide knowledge, experience, emotional social or practical help to each other [30]. Such support is essential since, the information that are generated through personal experiences are usually the most influential [2]. When it comes to coaching, peer support has the advantages of a low-cost intervention that reduces hospital stays, limited access to care [12] and generally has capability of

helping individuals to alter their own behaviour [23]. Another advantage of peer support is that it can be realised through various modes of interaction and involvement, in different settings and structures [12].

2.4 Summary

Health coaching is proven to be beneficial [38] for all stakeholders of public health systems [30] and thus should be widely adopted for the treatment and education of chronic disease patients. Nonetheless, the variety of intervention types and approaches has, thus far, prevented the foundation of standard rules and guidelines for health coaching. Consequently, health coaching cannot uniformly be applied into routine health care [26].

The key for a positive outcome in a coaching process, which is independent of the intervention used [1], is the interpersonal relationship between patient and physician. Hence, even though artificial intelligence(AI) agents can be used for health coaching and provide a sense of social interaction [15, 28] they might never totally replace this interpersonal interaction. However, due to the increasing number of patients, physicians and clinicians cannot solely deal with all the patients. Therefore, a hybrid model of coaching system, incorporating peer support, artificial and human coaches, should be researched.

In order to set the foundations for such an intervention and keep a balance between quality and cost, the level of involvement of AI and human coaching should be further examined. Therefore, the following research questions need to be answered:

- How and what can be artificialised in the practice of heath coaching for chronic diseases?
- How artificialisation will ameliorate the well-being of patients with chronic diseases and empower them?
- How a potential incorporation of AI coaches will abate health-care systems costs, while retaining each patient's singularity?

3 Proposed Model

The proposed approach of this paper comprises of a conceptual large scale health coaching intervention model that can be applied by any public health body. The model could be implemented as a cross-platform application including some or all of web based interventions, wearable technologies and monitors, mobile applications and a combination of sensed and self-reported

information inputted by the patient, their friends, family or peers and associated health and social care professionals.

The model unburdens clinicians and physicians and distributes workload to AI coaches. By downscaling the professionals' involvement to the coaching process and increasing peer support networks and AI coaching, an instant cost reduction will be made. These coaches will not totally cease the interpersonal interaction but instead reduce it to necessary intervention. The implementation of such a model requires some resources in order to be developed, but the long term saving should cover the investment.

3.1 Levels

3.1.1 Peer Support Level

On the lower coaching level, people are forming social networks and peer support groups in order to support and inform each other. Hence, empathy, compassion and other similar feelings can be created and the patients can feel part of a team and motivate each other. The application should help the patients create, administrate and edit peer support forums and blogs. Moreover, it should exploit the power and widespread of social media and incorporate some of their features (eg likes, sharing), in order to enhance the peer support process. Hence, patients can save time and share experiences more easily.

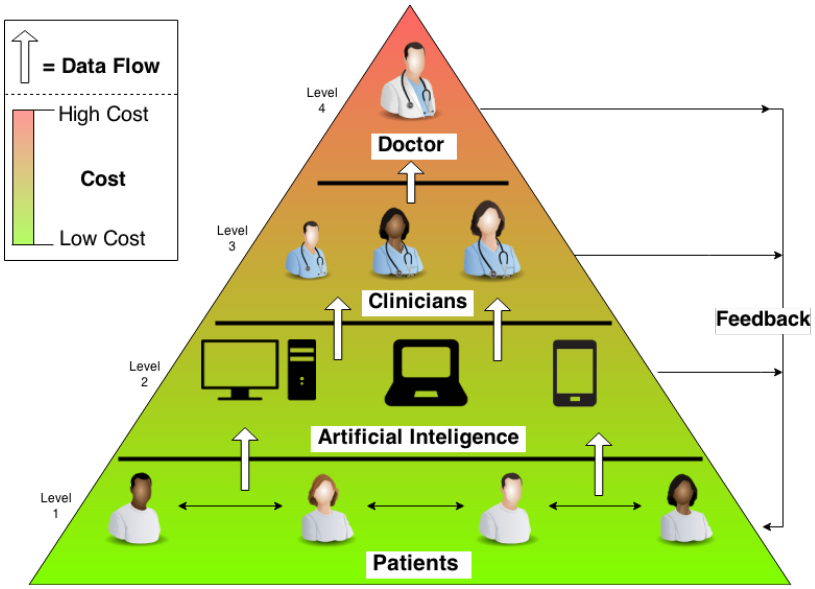


Figure 1. A multi-level hierarchical architecture that incorporates human and AI coaching. The higher in the pyramidal structure the fewer resources needed.

3.1.2 Artificial intelligence level

On the second level, coaching is taking place, where computers and smartphones are monitoring and guiding patients at frequent intervals. The application should be able to collect and use data provided by monitoring devices. Guidance by the AI coaches is limited to advice about everyday issues. This advice will be tailored to each patient. This level can also convey raw monitoring or processed data about patients' progress to the higher levels of the architecture.

3.1.3 Clinicians' level

On the subsequent higher level, clinicians (nurses, social workers, medical assistants (MAs), community health workers or health educators [6, 20]) will be monitoring the overall progress of the patients' condition. The system can provide treatment suggestions to the clinicians for each the patient. The feedback from the clinicians does not need to be delivered in person but can be communicated through the application. Hence, an inbox-like message delivery can inform the patient about something that the clinician has observed in his case and give suggestions, praise, provide reassurance and reinforcement in order to alter behaviour and further motivate the patient [35].

3.1.4 Attending physicians' level

The last level is consisted by the patient's attending physician, an interpersonal relation of trust [36], which in some cases lasts for even years. This relationship and the authority of the physician (as a professional) [5] is proposed to stay as it currently is. Hence, the face to face interaction and coaching, does not need to be replaced by computer mediated interaction. The physician can use the system to monitor the everyday progress of the patient and explain to them how the specific outcomes impact his total condition. The application can gather all the data and produce reports and visualise data. Because of the previous' levels contribution to the coaching process the physician can intervene fewer times (in bigger intervals) and thus, gain time to deal with more patients.

4 Conclusions

The present paper proposed a holistic and multi-level conceptual model for digital health coaching. The aim of the model is not the replacement of human presence by computers, but a coaching strategy that will enable, assist,

promote interaction and help automate (where needed and possible) resource consuming processes.

The proposed model will ultimately enhance the management of chronic diseases, increase patient empowerment and decrease costs by redistributing the available resources more efficiently, across the whole eco-system. Therefore, it can potentially become a new standard for integrated health coaching interventions.

Acknowledgements

The authors would like to thank Mr Jamie Stevenson for his valuable help.

References

[1] Balint, E. The possibilities of patient-centered medicine. The Journal of the Royal College of General Practitioners. vol. 17. 1969. p 269.

[2] Bandura, A. Social foundations of thought and action. Englewood Cliffs, NJ. 1986.

[3] Bennett, H. D., Coleman, E. A., Parry, C., Bodenheimer, T., and Chen, E. H. Health coaching for patients with chronic illness. Fam Pract Manag. vol. 17. 2010. pp 24-29.

[4] Bodenheimer, T. A 63-year-old man with multiple cardiovascular risk factors and poor adherence to treatment plans. Jama. vol. 298. 2007. pp 2048-2055.

[5] Brody, H. The healer's power. Yale University Press. 1993.

[6] Brownson, C. A., and Heisler, M. The role of peer support in diabetes care and self-management. The Patient: Patient-Centered Outcomes Research. vol. 2. 2009. pp 5-17.

[7] Bruegel, R. B. Patient empowerment--a trend that matters. Journal of AHIMA/American Health Information Management Association. vol. 69. 1998. pp 30-33; quiz 35-36.

[8] Carver, C. S., and Scheier, M. F. Attention and self-regulation. Springer. 1981.

[9] Cialdini, R. B. Influence: The psychology of persuasion. New York, Morrow. 1993.

[10] Consolvo, S., Landay, J. A., and McDonald, D. W. Designing for behavior change in everyday life. Computer. 2009. pp 86-89.

[11] Demiris, G., Afrin, L. B., Speedie, S., Courtney, K. L., Sondhi, M., Vimarlund, V., Lovis, C., Goossen, W., and Lynch, C. Patient-centered applications: use of information technology to promote disease management and wellness. A white paper by the AMIA knowledge in motion working group. Journal of the American Medical Informatics Association. vol. 15. 2008. pp 8-13.

[12] Dennis, C.-L. Peer support within a health care context: a concept analysis. International journal of nursing studies. vol. 40. 2003. pp 321-332.

[13] Eakin, E. G., Lawler, S. P., Vandelanotte, C., and Owen, N. Telephone interventions for physical activity and dietary behavior change: a systematic review. American journal of preventive medicine. vol. 32. 2007. pp 419-434.

[14] Fogg, B. J. Persuasive technology: using computers to change what we think and do. Ubiquity. vol. 2002. 2002. p 5.

[15] Gabriele, J. M., Carpenter, B. D., Tate, D. F., and Fisher, E. B. Directive and nondirective e-coach support for weight loss in overweight adults. Annals of Behavioral Medicine. vol. 41. 2011. pp 252-263.

[16] Golubic, M. Lifestyle Choices: Root Causes of Chronic Diseases. Cleveland Clinic, Electronically. 2013. Date accessed: 29/05/2015. Available at: https://my.clevelandclinic.org/health/transcripts/1444_lifestyle-choices-root-causes-of-chronic-diseases.

[17] Grimes, A., and Grinter, R. E. (2007) Designing persuasion: Health technology for low-income African American communities, In *Persuasive Technology*, pp 24-35, Springer.

[18] Guldemond, N., and Hercheui, M. D. (2012) Technology and Care for Patients with Chronic Conditions: The Chronic Care Model as a Framework for the Integration of ICT, In *ICT Critical Infrastructures and Society*, pp 123-133, Springer.

[19] Harjumaa, M., and Oinas-Kukkonen, H. (2007) Persuasion theories and IT design, In *Persuasive Technology*, pp 311-314, Springer.

[20] Holland, S. K., Greenberg, J., Tidwell, L., and Newcomer, R. Preventing Disability Through Community-Based Health Coaching. Journal of the American Geriatrics Society. vol. 51. 2003. pp 265-269.

[21] Hovland, C. I., Janis, I. L., and Kelley, H. H. Communication and persuasion; psychological studies of opinion change. 1953.

[22] Huffman, M. Health Coaching: A New and Exciting Technique to Enhance Patient Self-Management and Improve Outcomes. Home healthcare nurse. vol. 25. 2007. pp 271-274.

[23] Joseph, D. H., Griffin, M., Hall, R. F., and Sullivan, E. D. Peer coaching: an intervention for individuals struggling with diabetes. The Diabetes Educator. vol. 27. 2001. pp 703-710.

[24] Kennedy, C. M., Powell, J., Payne, T. H., Ainsworth, J., Boyd, A., and Buchan, I. Active assistance technology for health-related behavior change: an interdisciplinary review. Journal of medical Internet research. vol. 14. 2012.

[25] Lewis, F. M., and Zahlis, E. H. The nurse as coach: a conceptual framework for clinical practice. Oncology nursing forum. vol. 24. 1996. pp 1695-1702.

[26] Lindner, H., Menzies, D., Kelly, J., Taylor, S., and Shearer, M. Coaching for behaviour change in chronic disease: a review of the literature and the implications for coaching as a self-management intervention. Australian Journal of Primary Health. vol. 9. 2003. pp 177-185.

[27] Miller, G. On being persuaded: Some basic distinctions. Sage Publications. 2002. pp. 3-16.

[28] Nass, C., Steuer, J., and Tauber, E. R. Computers are social actors. Proceedings of the SIGCHI conference on Human factors in computing systems. 1994. pp 72-78.

[29] Newman, P., and McDowell, A. Health Coaching :Conversations for Better Care Concepts, skills and application for the East of England Programme. NHS, Electronically. 2014. Date accessed: 29/05/2015. Available at: https://www.eoeleadership.nhs.uk/downloadFile.php?doc_url=1390737952_Lkjk_health_coaching_and_our_approach_in_east_of_englan.pdf&area_id=73.

[30] NHS. Health Coaching for Behaviour Change Programmes - Better conversations, better care. NHS. 2015. Date accessed: 29/05/2015. Available at: https://www.eoeleadership.nhs.uk/page.php?page_id=490.

[31] Østbye, T., Yarnall, K. S., Krause, K. M., Pollak, K. I., Gradison, M., and Michener, J. L. Is there time for management of patients with chronic diseases in primary care? The Annals of Family Medicine. vol. 3. 2005. pp 209-214.

[32] Palmer, S., Tubbs, I., and Whybrow, A. Health coaching to facilitate the promotion of healthy behaviour and achievement of health-related goals. International Journal of Health Promotion and Education. vol. 41. 2003. pp 91-93.

[33] Porter, S., and Weitzel, K. Remote Coaching. The Oxford Group, Electronically. 2014. Date accessed: 29/05/2015. Available at: http://www.oxford-group.com/pdf/Remote%20Coaching.pdf.

[34] Richardson, C. R., Buis, L. R., Janney, A. W., Goodrich, D. E., Sen, A., Hess, M. L., Mehari, K. S., Fortlage, L. A., Resnick, P. J., and Zikmund-Fisher, B. J. An online community improves adherence in an internet-mediated walking program. Part 1: results of a randomized controlled trial. Journal of medical Internet research. vol. 12. 2010.

[35] Samarel, N., and Fawcett, J. Enhancing adaptation to breast cancer: the addition of coaching to support groups. Oncology Nursing Forum. vol. 19. 1992. pp 591-596.

[36] Thom, D. H., and Campbell, B. Patient-physician trust: an exploratory study. The Journal of family practice. vol. 44. 1997. pp 169-176.

[37] UK Government - Department of Health. Ten things you need to know about long term conditions. UK Government, Electronically. 2012. Date accessed: 29/05/2015. Available at: http:/www.dh.gov.uk/en/Healthcare/Longtermconditions/tenthingsyouneedtoknow/index.htm.

[38] Watson, A., Bickmore, T., Cange, A., Kulshreshtha, A., and Kvedar, J. An internet-based virtual coach to promote physical activity adherence in overweight adults: randomized controlled trial. Journal of medical Internet research. vol. 14. 2012. p e1.

[39] Wolever, R. Q., Caldwell, K. L., Wakefield, J. P., Little, K. J., Gresko, J., Shaw, A., Duda, L. V., Kosey, J. M., and Gaudet, T. Integrative health coaching: an organizational case study. EXPLORE: The Journal of Science and Healing. vol. 7. 2011. pp 30-36.

[40] Wolever, R. Q., and Eisenberg, D. M. What Is Health Coaching Anyway?: Standards Needed to Enable Rigorous Research: Comment on "Evaluation of a Behavior Support Intervention for Patients With Poorly Controlled Diabetes". Archives of internal medicine. vol. 171. 2011. pp 2017-2018.

[41] Yach, D., Hawkes, C., Gould, C. L., and Hofman, K. J. The global burden of chronic diseases: overcoming impediments to prevention and control. Jama. vol. 291. 2004. pp 2616-2622.

Rome Wasn't Reached in a Day: How to Motivate Patients to Keep Walking?

Bert Vandenberghe[1], Jasper Vanhoof[2], Fabienne Dobbels[2], David Geerts[1]
[1]Centre for User Experience Research, iMinds – KU Leuven, Leuven, Belgium
[2]Centre for Health Services and Nursing Research, KU Leuven, Leuven, Belgium
[1]{firstname.lastname}@soc.kuleuven.be
[2]{firstname.lastname}@med.kuleuven.be

Abstract. We have to sit less and take more steps. This message is important for us all, and especially for patients with chronic diseases. As solutions to motivate us and to boost our physical activity make use of gamification, and by doing so focus on the fit and healthy, we are wondering how we could do the same for people who see their abilities decrease (e.g. patients with chronic diseases). We worked on a solution for transplant patients that highlights the cumulative effect of daily activity. During the design and evaluation of this solution, we were confronted with a number of questions on this topic of keeping patients motivated. During the workshop, we would like to discuss how we could keep patients motivated, and how we could implement a social component to let patients interact with their environment, without getting demotivated.

Keywords: Patient, Physical activity, Self-management, Step counter.

1 Introduction

Technology enables us to remain seated throughout the day for our own comfort. While living life from the couch can feel like the ultimate sophistication of mankind, it is a bad habit with profound effects on our health [2]. On average 25% of the population is not active enough [4], and we can expect that this number is even higher for people with decreased abilities, e.g. patients. However, a healthy lifestyle that includes sufficient physical activity would certainly benefit the condition of the patients.

Existing solutions to keep us physically active are based on known behaviour change models (e.g. Fogg Behavior Model [1]). Activity trackers and accompanying dashboards raise awareness, let users set goals and achieve them, and have a social component to show off progress and facilitate competition between friends. Gamification, where game elements are used to

facilitate behaviour change, is applied widely. But we argue that these tend to focus on the fit, who might be able to increase their step count day by day, in order to earn badges and to become the leader among their friends.

For patients who see their abilities decrease, these approaches could be counterproductive. As multiple transplant patients told us, their disease is not a game. Being confronted with your limitations as a patient can be very demotivating. The use of step counters to support patients to increase physical activity has proven to be successful in other studies, e.g. for diabetes patients [3]. So we worked on a solution to keep transplant patients motivated by highlighting the cumulative effect of walking every day, because every step counts. Along the road, some remaining questions were raised which we would like to discuss further.

2 About the PICASSO-Tx project

The PICASSO-Tx project wants to investigate the influence of preference on the outcome of interventions that focus on self-management assisted by technology. Therefore, a technology solution was designed that supports the self-management of physical activity, medication adherence, and a healthy diet. Participants receive a Fitbit One activity tracker to register steps, Fitbit Aria weight scale to register weight, and an Aardex pillbox that registers the time of opening the pillbox. We wanted to have one low-barrier point of access to all data, in Dutch (as most of our participants don't speak English), and that meets the existing regulations regarding security and privacy of (big) data in healthcare. Therefore we decided to design this web application ourselves, as most commercial solutions don't meet our requirements. After all, it was utterly important that participants would not reject the technology for user experience issues that could have been prevented.

During one year, we followed a user-centred design approach in intense collaboration with researchers within the medical domain. We performed a diary study with twenty participants to map the current needs and existing solutions for the self-management of physical activity, medication adherence, and healthy eating. Based on this knowledge, we designed the interface of the web application through iterative prototyping. We invited patients to our lab to gather feedback on a prototype and then repeated this exercise. The final design of the PICASSO-Tx web application was finished by a graphical design team and was then validated in usability tests. To conclude the design phase, we performed a field test where twenty participants received the activity tracker, weight scale, pillbox, and access to the web application for

two weeks. We evaluated the solution using surveys and a semi-structured interview based on the critical incident technique.

In the next paragraphs, we will discuss our experiences with physical activity and the activity tracker throughout the design phase of the PICASSO-Tx project. We argue that the standard approach used in games for health is not suitable for our target group. We list a number of remaining questions that were raised during the design and evaluation of the PICASSO-Tx web application.

3 A pilgrimage that keeps patients walking

The PICASSO-Tx application offers a very limited simple view on the patients activity. The application allows patients to view their daily amount of steps for the past seven days. As the display of the activity tracker also displays the total number of steps, people were familiar with this daily step count. Our field test showed that people do look at this number throughout the day. Other details and stats (e.g. number of floors, number of active minutes) are not displayed to keep the interface as simple as possible. Furthermore, the stats functionality does not allow people to look at activity history for more than seven days at once, to avoid confronting the patients with a potential strong decrease in abilities.

Patients also have a goal, which they can try to reach in one week's time. The physicians set this goal, based on a baseline measurement. It can be adjusted according to the changing abilities of the patient during a face-to-face meeting with the physician, once every few months. If the goal is reached for at least five of the last seven days, it is raised with a few hundred steps – this leaves the patient two extra days to buffer a bad day or a visit to the hospital with less physical activity (something which happens quite frequently). If the patient does not reach the goal, a motivational message is displayed.

For the long term, we implemented a pilgrimage that highlights the cumulative effect of daily activity in order to motivate patients. If you take 150 steps per day, you walk a marathon in a year time. This idea was implemented on the PICASSO-Tx web application. Patients would virtually walk to e.g. Santiago de Compostela, a well-known Christian pilgrimage destination. For patients without Catholic background (or religion in general), other destinations could be selected that leads to the required motivation. The field test showed that this aspect of the feedback was especially well received by our participants. During our final interview, people were proud they reached France in two weeks time. Some of the participants reported that walking to Compostela feels like a dream, but their condition would never

allow them to do this. Even though our pilgrimage is virtual, it does appeal to our target audience.

4 The Road Ahead: Some Open Questions

Our solution will now be tested in a randomized controlled trial, to see if patients find motivation in the long term. We are interested in how seasonal changes will affect the physical activity of our patients. And for this group of patients, a (temporary) relapse in abilities is common. The same mechanisms could be at play here as well. Our field study showed that patients would like to involve their environment or family in this type of solutions. We did not implement this, as we did not want the competition between patients and their able-bodied environment to demotivate the patient. We also did not implement interaction between patients, as most patients stated during our diary study they don't like to be treated as patient. Still, we are wondering how social interaction between patients and their environment can be organized in an optimal way. During this workshop, we would like to discuss how to keep people motivated, even though their abilities decrease. And, how can we implement a social component for this group of people?

Acknowledgements

The research described is part of the PICASSO-Tx project, funded by Bijzonder Onderzoeksfonds KU Leuven (OT project OT13/101). We would like to thank Sabina De Geest, Thierry Troosters, Patrick De Mazière, Pieter Philippaerts, Dirk Kuypers, Frederik Nevens, Johan Vanhaecke, Geert Verleden, Lieven Dupont, Diethard Monbaliu, Jacques Pirenne, Dirk Van Raemdonck, Johan Van Cleemput, Robin Vos, Annette De Vito Dabbs, John Hughes, and Bernard Vrijens.

References

[1] Fogg, BJ 2009 A Behavior Model for Persuasive Design. In Proceedings of the 4th International Conference on Persuasive Technology Pp. 40:1–40:7. Persuasive '09. New York, NY, USA: ACM.

[2] Proper, Karin I., Amika S. Singh, Willem van Mechelen, and Mai J. M. Chinapaw 2011Sedentary Behaviors and Health Outcomes among Adults: A Systematic Review of Prospective Studies. American Journal of Preventive Medicine 40(2): 174–182.

[3] Qiu, Shanhu, Xue Cai, Xiang Chen, Bingquan Yang, and Zilin Sun 2014Step Counter Use in Type 2 Diabetes: A Meta-Analysis of Randomized Controlled Trials. BMC Medicine 12: 36.

[4] WHO | Physical Activity. WHO. http://www.who.int/mediacentre/factsheets/fs385/en/, accessed June 3, 2015.

Workshop:
The Landscape of UX Requirements Practices

UX Requirements to Public Systems for All: Formalisation or Innovation

Jane Billestrup, Anders Bruun and Jan Stage
Aalborg University
Department of Computer Science
Selma Lagerlöfs Vej 300, DK-9220 Aalborg East, Denmark
{jane,bruun,jans}@cs.aau.dk

Abstract. Many countries are developing e-government applications for digitalisation of the interaction between citizens and government administrations. To be successful, such applications must be usable and provide good user experience for all. In Denmark, e-government applications have traditionally been developed through a contract-based approach; but the experience has been quite negative, in particular in terms of user experience and innovation. To increase the user experience and provide a broader range of innovative solutions, the Danish government and the organization of the municipalities have produced guidelines and material for a more user-centred development process for the ongoing digitalisation of local government services. We present the guidelines and material together with the findings from case studies in four IT companies, where we have interviewed employees and conducted redesign workshops. Our findings indicate that the material and guidelines are a step forward, but they are too general and have failed to ensure a reasonable level of usability and user experience.

1 Introduction

An increasing number of e-government applications are created to reduce or augment face-to-face contact between citizens and employees of municipalities. The success of such applications depend critically on usability and user experience. Empirical studies have found that if an e-government website has a high degree of usability, citizens are more likely to accept the website, and keep using it [11], [12].

Countries like the United States and the United Kingdom are considering the importance of usability when designing interfaces for e-government [14]. Nevertheless, Wangpipatwong et al. found that e-government websites in several countries lack usability due to poor design and non-employment of user-centred design methodologies [18]. In South Africa, guidelines for designing e-government websites do exist but are generally not being applied by the web designers of the South African Provincial Government [15], [16], [17]. It is essential that citizens view e-government websites as both credible and reliable and have a high level of usability and user experience. One approach to accomplish that is to involve the end-users. The goal in applying

user-centred design is that the system serves the user and that their needs influence the interface design [13].

The Danish municipalities are in the middle of a digitalisation process with the end-goal that by the end of 2015, 80% of the interaction between citizens and municipality employees, that was previously based on paper forms will be handled digitally [1]. An example of these is an application for a new driver's license. Denmark has a population of 5.6 mio. people and is divided into 98 municipalities which serve as the single point of contact for citizens in regards to the public sector [29]. The digitalisation effort in this domain is in line with the European Commission's initiative "Digital Agenda for Europe" that defines a set of actions for digitalisation of the European Union. Here, action number 64 is "Ensure the accessibility of public sector websites", and the aim is that the public sector websites for citizens should be fully accessible and usable for self-service by 2015 [2].

So far, development of e-government applications for Danish municipalities has generally employed a contract-based approach. With this, the development of a software system is based on a formal contract between a customer (usually a single municipality) and an IT company, where the contract includes a fixed specification of requirements to the application. The advantage of this approach is that there is little uncertainty about the application that will be delivered. However, there are numerous disadvantages, particularly for applications where the requirements are unclear or even changing over time. The contract-based approach typically involves posting of formal bid material (or call for tenders) that IT companies use for making their proposals. Then the bids are assessed, a single IT company is selected and a contract is signed. It has been argued that this approach implies that requirements that are not mentioned in the bid material and the contract are plainly ignored. The IT company that obtains the contract has no incentives to consider additional requirements that are not included in the contract; and often the contract has very limited focus on user interaction, usability and user experience because these aspects appear to be difficult to specify.

This has led some to argue in favour of formal user experience requirements that can be objectively verified. The motivation for this workshop states that "This problem of omission or poor formalization of UX requirements is limiting the success of projects in the public and private sectors."

It seems doubtful that a more formalised contract-based approach is viable, because in the e-government domain, the disadvantages of a contract-based approach are even greater than in the general case due to the nature of this domain [9]. Development of e-government applications involves a broad array

of different stakeholders, including citizens, public institutions such as municipalities, support organizations like an IT organization that is servicing a group of municipalities, IT companies that produce applications and third party purveyors that the public institutions use to provide services to the citizens. It has also been documented that user-centred design is particularly difficult to facilitate when a contract-based approach is employed for development of e-government applications. An important reason is that some of the stakeholders are difficult to involve in a contract-based approach [7, 8]. The most important of these is the group of prospective users.

Even if stronger formalisation may resolve some of the problems originating from limited focus on user experience, the lack of innovation will remain a key problem. When a contract is made between a single municipality and a single IT company, the individual municipality will not be able to choose between competing designs. Formalisation of requirements must be balanced against other factors, such as identifying user classes, introducing innovation, and ensuring consistency among products from the same IT company. For these reasons, the Danish government and the joint organisation of the municipalities in Denmark have decided on a different approach. A key aim is to make the design process more user-centred and to provide the municipalities with a range of e-government solutions developed by different IT companies [3].

This paper presents an empirical study of the user-centred approach that is being employed in the Danish digitalisation process. In the following section, we describe the material that has been developed to facilitate user-centred design in the development process and how the IT companies have been supported in their development of IT solutions. Then we present the method of our study of the IT companies. This is followed by a presentation of the findings of the study with focus on the way the guidelines and material was perceived by the IT companies. In the conclusion, we discuss our findings and experiences in relation to similar work.

2 Guidance Material and Supporting Activities

Denmark is in the process of digitalising a significant amount of the services that municipalities provide to citizen. The goal is that by the end of 2015, 80% of the forms that have previously been completed by citizens for the municipalities will be filled in and submitted digitally [1]. This strategy was set to be deployed in four waves. The first wave was deployed in December 2012 and the last wave in 2015. Each wave released a new set of digital applications. This study was conducted in 2013-14 focusing on the

development of applications for the second wave. The plans for these four waves are shown in Table 1 on the following page.

	Phase 1 December 2012	**Phase 2 December 2013**	**Phase 3 Ultimo 2014**	**Phase 4 Ultimo 2015**
Service areas	• Notification of address change • Application for day care • Application for registration in elementary school • Application for after school activities • Application for a health insurance card • Application for EU health insurance card • Book in nature • Payment for hunters' test • Application for admission • Payment of student loan	• Notification of emigration • Application for protection of name and address • Application for loan to pay real estate tax • Application for free day care • Application for free after school activities • Change of general practitioner • Application for aid to burial expenses • Reporting of rats • Application for assisting techno-logy • Application for subletting facilities or buildings • Application for marriage certificate • Application for passport • Application for a new drivers' license	• Garbage handling for citizen • Garbage handling for organizations • Application for construction work • Application for building permission • Application for loan for deposit for dwelling • Application for citizen registration • Services in roads and traffic area • Notification of digging or work on pipelines • Certificates for lodging • Application for parking permits	• Application for sole providers • Application for aid for maintenance • Application for personal supplement • Application for sickness benefit

Table 1. Phases of digitalisation of self-service areas in Danish municipalities [25].

Before the outset of the process, the government and the joint IT organisation of the municipalities in Denmark decided to employ a new approach. Instead of the traditional development process based on a contract with a fixed set of requirements, the municipalities' joint IT organisation developed guidance and the following material to support a user-centred approach:

- User Journey
- 24 Usability Criteria

The purpose of this material was to secure accessibility and keep a user-centred focus in the developed self-service applications. The joint IT organisation of the municipalities functioned in a supporting role during the development process. All interested IT companies could decide which specific services they wanted to develop. The services were produced and made available for all of the 98 municipalities in Denmark. The municipalities can buy individual solutions and are not bound by one self-service provider as they can choose freely between all developed applications in each area.

Figure 1. Selected drawings from the user journey for applying for assisting technology for handicapped or elderly [24]. The short texts are in Danish and describe how Rita got injured

some years ago. The doctor finds she needs an insole. Both interact with the system in order to apply for that. After approval, the insole is made.

2.1 User Journey

The user journeys can be described as a person in a use situation described in a scenario [23] using graphical illustrations. The user journey is a graphical illustrated story describing how a typical user will interact with the self-service solution. A user journey was created for each self-service solution in the second wave. The user journeys were made by the joint IT organization of the Danish municipalities. The user journey for a specific self-service was made well before the IT providers would start developing that solution. The user journeys were developed on the basis of meetings in a focus group including both citizens and case workers at the municipalities. This was done to ensure that the user journeys would reflect an actual real use situation and the users' needs. Once the content was defined, the drawings were made by a professional artist.

The purpose of the user journey was to give both municipalities and the IT self-service providers an understanding of when and how users could interact with each system, and to make sure that the end-users were kept in mind during the development process of the self-service solutions. A user journey was created for each of the specific self-service focus areas in the second wave. A total of ten different user journeys were developed. Six segments from a user journey can be seen in Figure 1 on the following page.

2.2 Usability Criteria

The usability criteria are a set of guidelines with the purpose of describing how the IT self service providers could ensure that their applications were usable for all citizens. The criteria were defined at a 2011 IT and Tele Administration workshop focusing on usability [26]. It is stated that the criteria were defined by experts. The criteria were divided into four main focus areas, each consisting of six sub-categories. All 24 usability criteria are shown in Table 2.

2.3 Supporting Activities

The joint IT organisation of the municipalities has taken several approaches to supporting the self-service providers. They have hosted meetings and workshops where all the self-service providers were invited, and they have had meetings on a regular basis with each self-service provider. Additionally, the

municipalities' joint IT organisation phoned the self-service providers on a regular basis asking for a status update and offering their help and services when needed.

	Language and text
1	Texts are short and precise without containing legalese or technical terms
2	Text should be action-oriented and help the citizen
3	The citizen is informed of which documents to attach before filling out the form
4	The citizen can access additional information if needed
5	If an error is made it should be made very clear what is wrong
6	Error messages should be in Danish
	Progress and flow
7	The form should be clear for the citizen
8	The extent of the form should be clear for the citizen
9	When filling out the form the citizen knows the progress made and how many steps are left
10	The receipt should be clear to see after finishing the form
11	The receipt should also be sent by email to the citizen
12	The next steps should be clear to the citizen after submitting
	Data and information
13	If login is required, NemLogin (National Danish Identity Service) should be used
14	Excising data should be reused as much as possible so a citizen should not give the same information more than once.
15	A summary is shown before sending the form
16	Sending a form should only be possible if all required information is present
17	The solution should validate the typed information as much as possible
18	The solution should adapt as much as possible during the flow
	Design and accessibility
19	It should be clear when filling out the form begins
20	There should be a clear distinction between positive and negative buttons, and the positioning should make sense
21	The authority behind the form should be clear
22	Navigating in the form can be done both using mouse and keyboard
23	The form can be filled out by the citizen without possessing special skills
24	The solution meet relevant accessibility criteria for self-service solutions

Table 2. The 24 Usability Criteria

3 Method

This study was conducted as an empirical case study. The data was collected in 2013-2014. To get an overview of the development of self-service applications for the second wave, we initially interviewed one Project Manager from from each of the 11 IT companies identified as developing self-service solutions for this wave [30]. Thus the aim was to cover all 11 IT providers on an overall level.

Based on these initial interviews, we selected a self-service area and four IT-providers that we would focus on. As self-service area, we chose application for assistive technologies for handicapped and elderly; an example of this is application for a hearing aid. As IT-providers, we chose, form the total pool of the 11 IT providers, four companies that were developing a self-service solution for this service area. We interviewed a total of 14 people working in these four organisations. In addition to the interviews, we had one half-day meeting and one workshop with each of the four IT companies. The aim of these activities was to study the development process of the solutions for this self-service area in more into detail.

3.1 Participants

Four IT companies participated in this study. The application for assistive technologies for the handicapped or elderly was chosen because there were four self-service providers developing this solution which varied in maturity level. Two companies had an existing solution already in use by the municipalities that they were developing further, while the other two were new in this self-service area and were developing brand new solutions. The four organisations were divided as shown in Table 3.

	Immature organisation	Mature organisation
New self-service solution	Organisation A	Organisation B
Optimisation of existing self-service solution	Organisation C	Organisation D

Table 3. The IT companies chosen for this study.

The differentiation between mature and immature organisations was made in regards to developing self-service solutions. Of the new self-service providers, one was brand new in regards to self-service solutions. The other company was new in regards to the application for assistive technologies for

the handicapped or elderly but had developed several other self-service applications in Denmark.

3.2 Preparations

When starting this study we had one meeting with each of the four IT companies.

These meetings each lasted half a day. The project manager and the product owner were present, and in some of the IT companies a developer and a user experience designer were present as well.

At the meetings we were given a presentation of the development method used by the IT companies and how it was used in practice. We also received a demonstration of the self-service solution they were developing along with insights into how they worked with an on-site customer and their focus areas during the development process. At the end of the meeting we identified which people we wanted to interview as part of this study.

3.3 Procedure

To make certain all relevant people were interviewed for this study, we identified a set of relevant job functions that were perceived as important for the development process and which had extensive knowledge and different responsibilities in regard to the development process, knowing that some people might possess more than one of these job functions. The identified job functions were the following: Project Manager, User Experience Designer, User Interface Designer, Product Owner, Software Developer, and Market Segment Analyst.

We conducted between two and four interviews in each IT company, totalling to 14 interviews.

Three months after the first meeting we had a redesign workshop with each IT company. In that time period all interviews had been conducted and analysed. This meeting was conducted as a workshop in each company where the results from interviews were discussed and focus areas were identified. The entire preliminary conclusion from the interviews were discussed, processed, and modified in the workshops.

3.4 Data Collection

As part of this study four different methods were used for collecting data. We had one half-day meeting with each IT company. We conducted semi-structured qualitative interviews with two to four people involved in the

development process of the self-service solutions from each company, We completed a content analysis of relevant documents from both the municipalities' joint IT organisations and companies, and we hosted a re-design workshop with each of the four companies.

All interviews were conducted as semi-structured qualitative interviews as described by Kvale [4]. The interviews lasted between 25 and 59 minutes each.

The interviews established clarity in regards to the following:

- the interviewee's job function and level of experience
- the development process, including strengths and weaknesses
- the view of the user journey and usability criteria, including its strengths and weaknesses
- establishing whether the user journey and usability criteria were usable for the self-service providers
- missing elements in the existing materials and ways to improve this

After all interviews were conducted, the data was analysed in regard to the different perspectives of each interviewee and their job function in regard to the development process.

3.5 Data Analysis

Documents were gathered both from the municipalities' joint IT organisation and some of the IT companies. These were analysed and the results were used in correlation with the interviews. The interviews were transcribed and both interviews and documents were analysed using Dedoose (www.dedoose.com). All findings were added to a list that became the topics for the workshop discussions. All workshops were recorded. After the workshops were conducted, the recordings were transcribed.

The results from this study emerged in two steps. After the interviews were analysed, a list with our findings was created. This list contained all statements regarding the strengths and weaknesses described in relation to the user journey, usability criteria and self-service providers' communication with the joint IT organisation of the municipalities.

These identified weaknesses were discussed at four workshops, one with each participating IT company. These workshops led to a set of guidelines describing how to make the existing material more user-centred and which focus areas were currently not addressed in the existing material or supporting activities.

4 Findings

First we describe the findings from the conducted interviews, followed by suggestions for improving the user-centred approach.

4.1 Findings from the Interviews

Findings from the interviews are divided into three sub-sections describing perceived strengths and weaknesses of the user journeys, the usability criteria and the supporting activities. These findings identify the perceived strengths and weaknesses in regards to the development of the four self-service solutions and the companies' development process.

4.1.1 User-Centred Approach

The concept of involving User-Centred Design in the development process and creating user journeys was primarily described as a useful idea. The user journeys were generally described as neatly graphically created and helpful in regards to keeping focus on the end-user when designing the e-government applications. On the other hand none of the interviewees found the material to be a support in regards to developing self-service applications with a high degree of usability.

4.1.2 User Journeys

In regards to the user journey for application for assistive technologies for the handicapped or elderly, it was primarily used by the IT self-service providers in preliminary meetings with the municipalities as a tool for aligning expectations between the self-service providers and the municipalities.

“The user journey has been a strong tool for opening the dialogue with the municipalities.”

A few interviewees did describe some instances in which the user journey had set some expectations at the municipalities which the self-service providers then had to correct.

“Some municipalities thought we could deliver everything described in the user journey. They got quite disappointed when they realised we only deliver a small piece of the puzzle.”

The purpose of the user journeys was not communicated well, as some interviewees described that both they and the municipalities found it unclear whether the user journeys were to be perceived as a set of requirements or as a vision of how the citizens were expected to be interacting with the municipalities in a near or far future.

Most interviewees found the user journey useless in their analysis of the target user group for two reasons. The first is because the user journey only described one of many possible use situations and the second because the user journey was released too late in the process for them to use it in their preliminary analysis of the target user group and system requirements.

4.1.3 Usability Criteria

The 24 usability criteria are described as a mix of technical requirements and guidelines such as what kind of language to use in the self-service solutions. Several interviewees described how the interpretation of the criteria has been difficult at times, and several interviewees found themselves interpreting the criteria differently than intended by the municipalities' joint IT organisation. Several interviewees stated that this slowed their development process as this wrongful interpretation was not discovered until a later time, causing them to have to go back and restructure in order to meet these requirements. This was described as frustrating and several stated that they felt the usability criteria should have been described into more detail.

"I think they could have done a better job making the criteria understandable and user friendly for the self-service providers."

Though this material was called usability criteria, the interviewees responsible for the usability and user-experience design of the developed self-service solutions felt that the usability criteria did not ensure that the self-service solutions would actually become usable for all citizens. The interviewees expressed opinions that the criteria lacked focus regarding actual use and usability.

4.1.4 Supporting Activities

Though most interviewees were positive in regard to the support they received from the municipalities' joint IT organisation, they also found room for improvement, especially in regards to release of time schedules and the supporting material. The interviewees also mentioned that the joint IT organisation should put more effort into making sure that supporting solutions, such as the power of attorney, which should be implemented into the new solutions, were released on time. Several stated that they had a very tight deadline to develop and implement the self-service solution, but they were delayed because they had to wait for others to finish the specific parts the self-service providers were required to implement into their systems. The self-service providers felt that the joint IT organisation of the municipalities should put more focus into making sure these portions were finished on time.

During this study we found several misunderstandings about the communication between the municipalities' joint IT organisation and the self-service providers. For example, all self-service providers thought that the usability criteria were mandatory to implement, leaving them struggling to understand and implement these criteria into their solution, but we found later that the usability criteria were only intended as guidelines. Several times we had one understanding from all self-service providers, but later learned that the joint IT organisation of the municipalities had a very different understanding. For the user journey, we found that some self-service providers thought it was meant as a set of requirements that the developed self-service application had to meet, but the intention from the joint IT organisation of the municipalities was that the user journeys were meant as inspiration to help keep focus on the end-user.

Several interviewees described having trouble finding the documents or supporting materials they needed from the joint IT organisation of the municipalities. Even though the needed materials should be readily accessible on a website, several interviewees described that they had difficulty finding what they needed on this website. The website was mainly described as confusing and the search function was not helpful in regards to this matter.

4.2 Suggestions for improvement

After the workshops we created a set of guidelines for strengthening the focus on user-centred design and enhancing the communication between the municipalities' joint IT organisation and the self-service providers. These guidelines were based on the discussions from the workshops. Each workshop processed the same topics, but the workshop with the second self-service provider was also based on the results from the first workshop, and so forth. It would have been preferable to host a single workshop including all four self-service providers, but as they are competitors and based in different parts of the country, it was not a feasible solution. Overall, four foci areas were identified that needed to be optimized: Clearer communication, widening the focus to include all parts of the system, and not just the front-end, strengthen the involvement of all stakeholders, creating more user-centred material, and implementing a user-centred focus. These five focus areas will be elaborated in the following section.

4.2.1 Clearer Communication

Lack of communication has been an issue. This has been less of an issue in the day-to-day communication, but more problematic in communicating the

purpose and intentions behind initiatives like the user journey and the usability criteria. The participants described feeling frustrated and confused from time to time. They also described employees at the municipalities feeling the same way.

It is important that the municipalities are part of the initiatives as they are the ones the joint IT organisation of the municipalities is representing. The case workers at the municipalities need to know the intentions behind the materials provided by the municipalities' joint IT organisation and how they will be able to use the materials to its full potential.

4.2.2 Widening the Focus

At the workshops it was made clear by the participants that the process lacked a sense of the system as a whole. It was described that the focus was primarily on the citizens' solutions at the front-end, but that this should go hand in hand with prioritizing the back-end as this will help optimize the flow of the whole process instead of creating two different systems that will de-optimize the work-flow.

As the focus is purely on the applications for the citizens, the system used by the case workers was not prioritized at all. Given that the aim was to save money in regards to the time that caseworkers use, this is a problem. Instead, the self-service solutions should be seen as one whole solution focusing on usability and efficiency in regards to both citizens and caseworkers.

4.2.3 Strengthen the Involvement of All Stakeholders

At the workshops it was described as very important to involve all stakeholders before developing materials supporting user-centred design. Stakeholders were divided into four different categories: citizens, municipalities, third party providers and IT providers.

Citizens should be represented within the target user group and involved in order to acquire an understanding of their needs and abilities. Some suggested involving societies such as those for the elderly or handicapped. Others were reluctant about this as they felt it was not ideal to involve societies that could set demands without having any responsibilities of their own.

Municipalities should be represented as their work-flows and procedures are very different. The case workers at the municipalities can also help with focusing on the correct group of end-users. The municipalities, as the purchasers of the IT solutions, need to be represented in the process as they are the ones who have to be able to use the materials to their full extent when buying the IT solutions.

Third party providers can be doctors, undertakers or surgical appliance makers. Some self-service forms are in all or most cases filled out by a third party provider. For example, the application for acquiring aid for a funeral is always filled out by the undertaker and not the relatives.

IT providers should be involved as they are the ones who will have to use the developed materials in practice. Involving them at an early stage will give them an opportunity to comment and point out deficiencies at an early stage.

4.2.4 Creating More User-Centred Material

At the workshops, three important areas were identified: vision, clarification of user needs and technical requirements.

These areas are already present in the existing materials, but they are mixed as the user journey consists of both visions and user needs, and the usability criteria consist of user needs and technical requirements. We suggested that existing materials could be redesigned into three separate pieces.

The vision would describe which requirements could be set in the future and which goals the municipalities' joint IT organisation wish to achieve with the self-service solutions. The vision should be revised as requirements for technology changes but should always keep focus on the interest of both the end-users and municipalities in regard to work-flow. By doing this, both municipalities and self-service providers would be able to understand what goals, existing solutions, updates and new solutions are important.

The clarification of user needs should describe several different types of users from the target user group and which special needs should be taken into consideration. This could be a collection of Personas as described by Nielsen [6] and focusing on special needs and requirements in relation to handicaps, nationality and age, depending on the target user group. This would give the self-service providers a thorough analysis of the end-users and their needs, and save the self-service providers time and effort. They all described not having time or funding for conducting a major user study themselves. If the joint IT organisation of the municipalities did this thoroughly, it would ensure all user segments would be taken into consideration during the development process of the self-service solutions.

The technical requirements should be created as a check-list targeted towards the software developers. This list should describe server response times and for which Internet browsers to optimise the software solutions. This would help the software developers to know exactly which technical requirements the self-service solutions had met, and it would provide the employees at the municipalities with a check list they could use when deciding which self-service solution to acquire for each self-service area.

4.2.5 Implementing a User-Centred Focus

The above suggested redesigns of the materials cannot stand alone in regard to acquiring a more user-centred approach both in regards to the citizens' usage and optimizing the work-flow of the caseworkers. This needs to be supported by conducting usability evaluations on all self-service solutions including the work environments of the caseworkers. These evaluations should be conducted by independent usability experts so all self-service solutions are tested on the same basis. Then all IT solutions could be rated and benchmarked, or in other ways quantified, to make it clear for the municipalities whether or not an IT solution is user-centred and usable. This would ensure the self-service providers are focusing on creating usable systems. This recommendation was also suggested by several interviewees and discussed at the workshops.

Both a formative and a summative evaluation should be conducted. The formative evaluation should be conducted early in the process and could be conducted using a paper prototype, which would make it fairly inexpensive to change the design and fix problems very early in the design process.

All self-service solutions should be user-tested at the end of the process by conducting a usability evaluation with citizens from different user segments, and then benchmarked as described above. This would mean that all self-service providers would have to keep a user-centred focus during the development process, and it would help the municipalities to acquire usable self-service solutions without major usability problems.

4 Discussion

Previous research shows that usability and user-centred design are crucial for designing e-government services [11-18]. Our study shows that implementing a user-centred approach is on the right track and the user-centred initiatives described in this study appear to be interesting and innovative. Nevertheless, the level of maturity is still low. In South Africa, guidelines for designing e-government websites have been created but are not being applied by the web designers of the South African Provincial Government [15-18]. In this study we found that wanting to implement user-centred design is not the same as actually creating a user-centred design. Creating and implementing tools such as a user-journey and usability criteria is a step in the right direction, but it takes time and more than one attempt to create materials like these that will actually improve the usability of the end-system.

Several researchers have argued that traditional methods for user-centred design are difficult or impossible to employ in the development of e-government applications. The arguments relate to the size of these projects [7] and the diversity of the user group [8].

A report from OECD on the European development of e-government services states very clearly that the focus on technology has for years overshadowed the need for organisational, structural, and cultural changes in the public sector. Therefore, key challenges and prerequisites for building attractive, integrated, user-focused e-government services have been left unaddressed [10].

This is in line with our findings where we have seen that some of the user groups have not been involved in the development of the IT services. Even though there has been a general interest in focusing on the users, citizens in particular, the actual involvement has been very limited. It is interesting that this is emphasized consistently by several of the IT companies who argue that the citizens should be more directly involved.

Some researchers have presented ideas for overcoming the challenges of involving citizens in the development of e-government systems. One idea is to include citizens directly in groups or through representatives [8]. Citizens were included by the joint IT organisation of the municipalities as various user groups were consulted when the user journeys and usability criteria were defined. However, our findings show that it has not been successful or sufficient.

Another possibility is to combine participatory methods with methods for technology assessment that have been tried in practice, although this requires a group that can drive these activities [7]. So far, that has not been implemented in the Danish digitalization project. It has also been suggested to use early prototypes as a means for verifying that the user requirements are correct [9]. However, the viability of this idea has yet to be demonstrated in practice.

Iivari and Iivari examine user-centredness as a multidimensional concept along four aspects: as user focus, as work-centredness, as user participation, and as system personalization [5]. User focus reflects the traditional approach in user-centred design. Work focus is concerned with the work activities of the users. User participation is the active and direct involvement of user. Finally, system personalization indicates that the designed system can adapt or be adapted to the user during use. The aim of the Danish digitalization project has been to achieve a strong user focus, although it has only been partly successful. The other three forms can be used as inspiration for further development. Unless there is a basic move in this direction, the intended degree of user take-up is unlikely to be realized [10].

Enhancing usability and designing with a user-centred focus is not only important in regards to the citizens. In Denmark the strategy of digitalising citizens' self-services was conducted with the purpose of saving money. Bruun and Stage found that redesigning a citizens' self-service application for applying for a building project like a garage could decrease the time spent by the caseworker from an average of 53 minutes to 18½ minutes [27]. This shows that a user-centred focus is not only for the sake of the citizens but is a key aspect in regards to saving money on implementing E-government self-service solutions.

Focusing on both the front-end for the citizens and back-end for the caseworkers is important in regards to saving money on e-government self-service solutions. This means that it should also be a priority to develop a usable system in regards to the caseworkers. Another study has shown that a new system at a hospital for patient charts was not found to be more usable for the staff even after they had actually been using the system for a year than it was immediately after the system was deployed [28]. This means that usability problems do not go away just because employees are using a system daily. Thus caseworkers in the municipalities are spending more time than necessary on each e-government application, compared to a system that was designed with a focus on usability from the start.

5 Conclusion

We have presented findings from an empirical study of the approach that is being employed in the Danish digitalization process as well as how it is viewed by the IT companies. We have focused on the materials that have been developed to facilitate user-centred design in the development process and how the IT companies have been supported in their development of e-government self-service solutions. Our findings show that supporting others in designing user-centred applications, while well-intended is not straightforward. Wanting to create materials to help others design user-centred materials need to be designed very thoroughly and there need to be an understanding of both the end-users and the IT companies that are meant to use the material. The designers of the user-centred material need to understand all aspects of the development process and the end-users' needs. This is a challenge and should not be taken lightly if designing user-centred material that others are supposed to use in regards to understanding and designing for a target user group. Key points are that the material designed to support the IT companies in designing user-centred is very general and fail to ensure a reasonable level of usability. Instead, we have suggested some areas that could

be improved in regards to communication, which include focusing on the entire system and not just the user-interface in regards to the citizens, and involving more stakeholders in the creation process of user-centred materials. Additionally, we suggested new materials to develop regarding vision, clarification of user needs and technical requirements. Our suggestion is that these initiatives are backed with conducting usability evaluations of all self-service solutions. The idea is that by conducting these usability evaluations, the self-service providers have to keep focused on creating self-service solutions that are usable and without significant critical usability errors. If all self-service solutions are evaluated and benchmarked, it will make it much easier for the municipalities to choose the most usable solutions.

This paper is based on interviews and other qualitative methods that have been used to discover the opinions of four out of the eleven IT companies that were involved in the development of the digital services. We have selected them to reflect the variety of IT companies, but we cannot guarantee that they are entirely representative. The findings presented in this paper indicate avenues for future work. The most urgent is to evaluate the actual usability of the systems developed so far. It is also vital to experiment with techniques for involving citizens actively in a user-centred development process for e-government applications.

Acknowledgements

We would like to thank the IT companies and employees that participated in our questionnaire survey and the Infinit network for supporting the research.

References

[1] Organisation of the Municipalities in Denmark. Laws about mandatory digital self-service and digital mail has passed, December 2014. http://www.kl.dk/Administration-og-digitalisering/Lov-om-obligatorisk-digital-selvbetjening-og-digital-post-er-vedtaget-id105354/

[2] European Commission. Digital Agenda for Europe, October 2014. http://ec.europa.eu/digital-agenda/en/pillar-vi-enhancing-digital-literacy-skills-and-inclusion/action-64-ensure-accessibility-public

[3] Kombit. For self-service providers, January 2014. http://www.kombit.dk/ind hold/leverand%C3%B8rer

[4] Kvale, S. Interview, København: Hans Reitzel, 1997.

[5] Iivari, J., Iivari, N. Varieties of User-centredness. In Proceedings of the 39th Hawaii International Conference on System Sciences, (2006). p. 176a-176a

[6] Nielsen, L. Personas – User Focused Design", Human-Computer Interaction. Springer, 2012.
[7] Oostveen, A.-M., van den BesselaarI, P. User involvement in large-scale e-Government projects: Finding an effective combination of strategies and methods. Proceedings of the INTERACT 2005 workshop on User Involvement in e-Government development projects, 2005.
[8] Oppermann, R. User participation supported by usability engineering experts in e-Government projects. Proceedings of the INTERACT 2005 workshop on User Involvement in e-Government development projects, 2005.
[9] Skjetne, J. H. Prototypes and tender documentation. Proceedings of the INTERACT 2005 workshop on User Involvement in e-Government development projects, 2005.
[10] OECD Rethinking e-Government Services: User-centred Approaches, 2009.
[11] Clemmensen, T. & Katre, D.: Adapting e-gov usability evaluation to cultural contexts. Usability of e-government systems (pp. 331-344). Boston: Morgan Kaufmann, 2012.
[12] Huang, Z. & Benyoucef, M.: Usability and credibility of e-government websites. Government Information Quarterley (pp. 584-595), 2014.
[13] Norman, D. A. Cognitive Engineering. In D.A. Norman and S.W. Draper (eds) User centred Systems Design. Hillsdale, NJ: Lawrence Erlbaum Associates Inc., 1986.
[14] Soufi, B. & Maguire, M. Achieving usability within e-government websites illustrated by a case studiy evaluation. Proceedings of Human Interface, part II, pp. 777-784. Springer-Verlag, 2007.
[15] Yeratziotis, A. The impact of cultural context on web design for e-government in South Africa. Master's thesis, Nelson Mandela Metropolitan University, 2008.
[16] Korsten, H. & Bothma, T. J. D. Evaluating South African government web sites: methods, findings and recommendations (part 2). South African Journal of Information Systems, vol. 7, nr. 3, 2005.
[17] Pretorius, M. & Calitz, A. The South African user experience maturity status for website design in provencial governments. Proceedings of the 12th European Conference on eGovernment, Institute of Public Governance and Management ESADE, pp. 589-599, 2012.
[18] Wangpipatwong, S., Chutimaskul, W. and Papastratorn, B. Understanding citizens continuance intention to use e-government website: A composite view of technology acceptance model and computer self-efficiancy. The Electronic Journal of e-Government, vol. 6, nr. 1, pp. 55-64, 2008.
[19] Billestrup, J., and Stage, J. "E-government and the Digital Agenda for Europe." Design, User Experience, and Usability. User Experience Design for Diverse Interaction Platforms and Environments. Springer International Publishing pp. 71-80, 2014.

[20] Larusdottir, M.K., Haraldsdottir, O. and Mikkelsen, B. User involvement in Icelandic software industry. Proceedings of I-Used 2009 workshop at INTERACT 2009, pp. 51-52, 2009.
[21] Lárusdóttir, M.K., Cajander, Å. and Gulliksen, J. The big picture of UX is missing in scrum projects. Proceedings of the 2nd international workshop on the interplay between user experience evaluation and software development, in conjunction with the 7th Nordic conference on human-computer interaction, 2012.
[22] Yin, R. Case Study Research: Design and methods. CA, USA: Sage Publications, 1991.
[23] Nielsen, Lene. Engaging personas and narrative scenarios. Handelshøjskolen, 2004.
[24] Kombit, User Journey for applying for a marriage certificate, December 2014 http://www.kl.dk/ImageVaultFiles/id_60480/cf_202/Illu.PDF
[25] Organisation of the Municipalities in Denmark. Wave plan for mandatory digital self-service, October 2014.
[26] http://www.kl.dk/Om-KL/Opdateret-version-af-bolgeplanen-for-obligatorisk-digital-selvbetjening-id114939/?n=0&section=4652
[27] The government agency for digitalisation in Denmark. Criteria for usability in self-service solutions, January 2015.
[28] http://www.kl.dk/ImageVault/Images/id_56064/scope_0/ImageVaultHandler.aspx
[29] Bruun, A. & Stage, J. Supporting Diverse Users: Implementing Usability Improvements of an E-Government Data Entry Website (unpublished), 2014.
[30] Kjeldskov, J., Skov, M.B. and Stage, J. A longitudinal study of usability in health care: Does time heal?. International Journal of Medical Informatics 79.6 (pp. 135-143), 2010.
[31] The Danish ministry of health. Reform of the municipalities, January 2015.
http://www.sum.dk/Aktuelt/Publikationer/Publikationer_IN/~/media/Filer-Publikationer-IN/Kommunalreformen/2005/Kommunalreformen-kort-fortalt/kommunalreformen-kort-fortalt.ashx
[32] Billestrup J. and Stage, J. (2014) E-government and the digital agenda for Europe a study of the user involvement in the digitalisation of citizen services in Denmark. Proceedings of DUXU 2014. Lecture Notes in Computer Science, Volume 8518, pp. 71-80.

User Experience in Technically Complex and Safety Critical Work Domains: Two Case Studies in Aerospace and Aviation

[1]Dorrit Billman, [1]Jessica Lee, [2]Emilie Roth

[1] San José State University &
NASA Ames Research Center, Moffett Field, CA 94035-1000
dorrit.billman@nasa.gov
jessxlee@alumni.stanford.edu

[2]Roth Cognitive Engineering
2 Oliver Court, Menlo Park, CA 94025
emroth@mindspring.com

Abstract. In technically challenging and safety critical work, the primary objective in software design should be supporting execution of that work. Here, experienced usefulness may be the primary contributor to a good "user experience". We report two successful case studies assessing work-needs and designing software; both support planning and scheduling. Case 1 addresses software for planning the movement of the International Space Station (ISS). Case 2 addresses software for planning airlift operations. We compare the processes and representations in the two projects and offer a few suggestions about their benefits.

1 Technical Work Domains and User Experience

We report two case studies for planning and scheduling software in complex, technical work domains. Case 1 addresses software for planning the movement of the International Space Station (ISS). Case 2 addresses software for planning airlift operations. We reflect on commonalities and differences, and what processes and representations might contribute to software that supports user needs in work domains such as these.

In work domains, particularly safety critical and time-urgent work, a key driver of utility is whether workers can use technology to accomplish work goals. In well-designed work domains, while individual and institutional goals may differ, they do not fundamentally conflict. Providing well-designed Human-Technology systems, then, implies providing good support to enable workers to accomplish work goals in a satisfying and satisfactory manner. In this context, then, 'user experience' is primarily oriented to goals originating

in the work domain, and outward facing to that external and constrained work; this contrasts with contexts where 'user experience' is primarily oriented to goals originating in personal, individual, or discretionary preference, and thus facing inward toward the user. These contexts produce different forces for allocation of development resources, product design, and marketing strategy.

In our cases, "user experience" was primarily addressed in this work-centered way. Further, in our cases, work was primarily determined at an institutional rather than individual level. There can indeed be disparities between individual and institutional goals; for example, users are often not the voice that drives development or purchase decisions and the view of the decision makers may not accurately capture the work from the perspective of those carrying it out. In addition, the values or measures of remote decision makers may be different from those of users. For example, remote metrics may focus on factors that are more easily measured and "objective" such as time and money, or even margin of safety. Some aspects of effectiveness such as time and accuracy of completing selected tasks can be measured directly and objectively. Other aspects of effectiveness may be much more difficult to measure, such as ability to regulate workload, to recover in unusual situations, to coordinate with entities outside the defined work group, or to discover improved methods of accomplishing work. For these more complex and difficult to measure aspects of effectiveness, User ratings and satisfaction, of both experts and novices, may be a proxy for this type of hard-to-measure aspects of effectiveness. Thus, as has been widely noted (Bisantz & Roth, 2008; Holtzblatt, Wendell, & Wood, 2005), involvement of users in design is beneficial for many reasons ranging from the intrinsic satisfaction it may provide to increased uptake of the developed technology. In our cases, the primary role of users was (with others) to characterize the work to be done, and to provide input on the appropriateness of methods and designs for accomplishing that work.

In short, UE in the context of work, and particularly safety critical work, often is and should be, primarily a matter of effectiveness and safety. Involving subject matter experts to characterize the structure of work is required, and almost surely will include the users of the software being developed. It may also involve other experts and caches of expertise, such as engineers or technical documentation of automation being controlled and other engineered or natural systems in the work domain. Critical knowledge may be distributed across multiple roles, not just users, which was particularly true for our Case 2.

Each case relied on practices oriented to ensure that the systems enabled users to effectively accomplish the work. Neither of the cases reported here

relied only on formal requirements to ensure that the developed system was useful, usable, and satisfying. However, both produced documents which collectively guided design and development toward systems that effectively served the users needs. The two cases differed in detail of process and of the representations to support their process.

We present the two cases with particular attention to the role of requirements, documents, and representations. We provide a summary comparison of process and representations used. We offer a speculative analysis of what practices used here will more generally lead to useful and usable system, particularly in agile programming environments.

2 Case 1: Planning the Movements of the International Space Station

The movement of the International Space Station (ISS) is controlled by the Attitude Determination and Control Officer (ADCO) group in NASA Mission Control at Johnson Space Center, together with their Russian counterparts. The ADCO group both builds plans and executes from these plans. An ADCO operator who was trying to improve an ADCO planning tool contacted researchers in the NASA Ames Human Systems Integration (HSI) group for assistance. The AID (Automation-Integration Design) research group at Ames studied a small part of ADCO's work needs, based on limited observation and on analysis of the plans that were products of this group (D. Billman, Feary, Schreckenghost, & Sherry, 2015). This led to a small prototype based on technology built by the Human-Computer Interaction (HCI) group, also within Ames HSI. A laboratory evaluation of the prototype showed large improvements on the selected tasks (D. Billman, Arsintescu, Feary, Lee, & Tiwary, 2015), and this data provided evidence to ADCO managers arguing for funding a larger project. The Ames HCI group had previously built and deployed highly successful planning tools for NASA missions. ADCO was able to secure the services of the HCI group to develop new, production-level planning software. This case study began with initial good will of customers and a technical base in the development group.

2.1 Project context.

Designed to replace multiple legacy planning tools and processes, the ADCO Planning Exchange tool (APEX) incorporates workflow management capability in addition to traditional planning and scheduling functionality. APEX was developed at NASA Ames Research Center by the Human-

Computer Interaction (HCI) group, a group with extensive experience delivering planning and scheduling software for NASA missions. Using contextual inquiry, the team designed and developed the Scheduling and Planning Interface for Exploration (SPIFe), an integrated planning and scheduling toolkit (Aghevli, Bencomo, & McCurdy, 2011). APEX was built by heavily adapting SPIFe to the needs of ADCO users. The project had a two-year timeline with predetermined release milestones and time allocated for testing and transition.

2.2 Team composition.

The team within the HCI group responsible for delivering APEX consisted of one HCI project manager, one HCI product manager who performed user research and design, and a development team of five engineers. From the ADCO group, three end users, all experienced ADCOs at Johnson Space Center, were part of the team: one ADCO project manager, one ADCO project engineer, and one ADCO test engineer. Project managers were responsible for setting the overall timeline, scope, and budget of the project. The responsibility of the HCI-product manager included user research, design, and feature prioritization; this was purposefully intended with the belief that the same individual with deep user and workflow knowledge could most effectively design appropriate solutions and inform feature prioritization with developers. The HCI product manager worked closely together with the ADCO project engineer (who liaised with the broader ADCO team as needed) via weekly phone calls and emails to execute the project. The concentrated amount of knowledge between these two individuals facilitated decision making, communication, and informal requirements gathering.

2.3 Requirements gathering process and work representation.

The foundations of APEX were largely influenced by the integrated planning and scheduling toolkit (SPIFe), previously developed by the HCI group using contextual inquiry and ethnographic field observations. In order to adapt SPIFe to the needs of ADCO users, the product and project manager spent two weeks shadowing ADCOs during their shifts to develop an overall understanding of ADCO workflow and tools. ADCOs, aware of the observers, operated normally and pointed out processes of interest, usually involving software. Notes were recorded in a two-column spreadsheet, with an "observation" column documenting objective observations and an

"interpretation" column used for noting implications with respect to product design.

The project managers also provided an initial high-level product requirement document, drawing heavily on required ADCO functions (see Figure 1). For example, requirements included exporting documents with specified information, tracking changes in the plan, and previewing files before sending. Requirements were listed in terms of tool functionality (not design requests) and remained largely stable through the development process. APEX met all requirements, and also provided enhanced automation, visualization, and information organization that significantly improved user experience; the latter was informed by cycles of user observation, testing, and feedback.

	Outputs As-Flown TRTL in the following formats
	as-flown.trtl
	as-flown.trtl.html
	as-flown.trtl.plan
	As-Flown TRTLs will be modifiable to track events as they were flown
	Outputs TCFs per the format defined in OIP
Exchanges UAFs, ATLs, TCFs, and TRTLs to appropriate tools and IPs.	
	Send and receive UAFs to/from MMC via drop-box
	Send ATLs to ESA-ATV via drop-box
	Send and receive TCFs to/from SUDN via drop-box
	Send TRTLs to SACE and PLATO
Ops Timeline, ATL and TCF updates should be completed with minimal user interaction	
	Changes made to the Ops Timeline that effect an ACR time, will result in the creation of a revised ACR, and updated UAF file for delivery to Moscow.
	Revised ACRs will be made available for review prior to sending to Moscow.
	After revised ACRs are sent to Moscow, the ATL will be updated with the new ACRs.
	Changes made in an ACR time by Moscow will be updated in the Ops Timeline. User will be able to review the change prior to it being implemented in the Ops Timeline.
	After ACRs have been implemented in the Ops Timeline, received or approved ACRs will be created for delivery to Moscow. Following delivery to Moscow, the ATL will be updated with the new ACRs.

Figure 1. Part of the initial high-level product requirement document. Requirements were largely driven by institutional and functionality expectations; design specifications (how the requirement should appear and be implemented) was not included.

Due to differences in team location, user research and more detailed requirements gathering were performed remotely via phone and screen share where the ADCO project engineer explained the use cases and details of each requirement. After gaining a basic understanding of the requirements and workflow, the product manager created a document, including a summary of requirements and work representations, outline of use cases, along with several design proposals (see Figures 2 and 3). This document served as an effective communication tool and point of reference for all subsequent conversations relating to the requirement; the ADCO project engineer could provide concrete feedback, discuss advantages and disadvantages of each design, share it with other ADCOs, and point out missing edge cases. After iterations of discussion and design feedback via informal verbal and email

conversations, the document provided sufficient detail so that it could be passed on to the developer to implement. Document content varied depending on the requirement; some requirements were more process-oriented (Figure 2), while some requirements focused more on providing visual cues for decision-making (Figure 3).

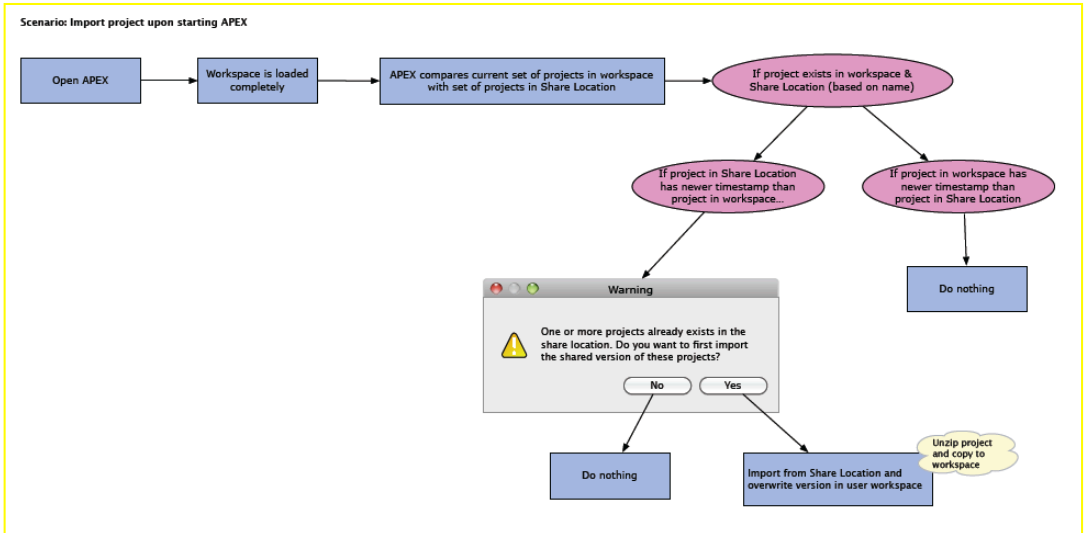


Figure 2. Part of the design document for the file sharing system. The document steps through the process of opening the software, checking against existing files, and providing the user with the option to overwrite.

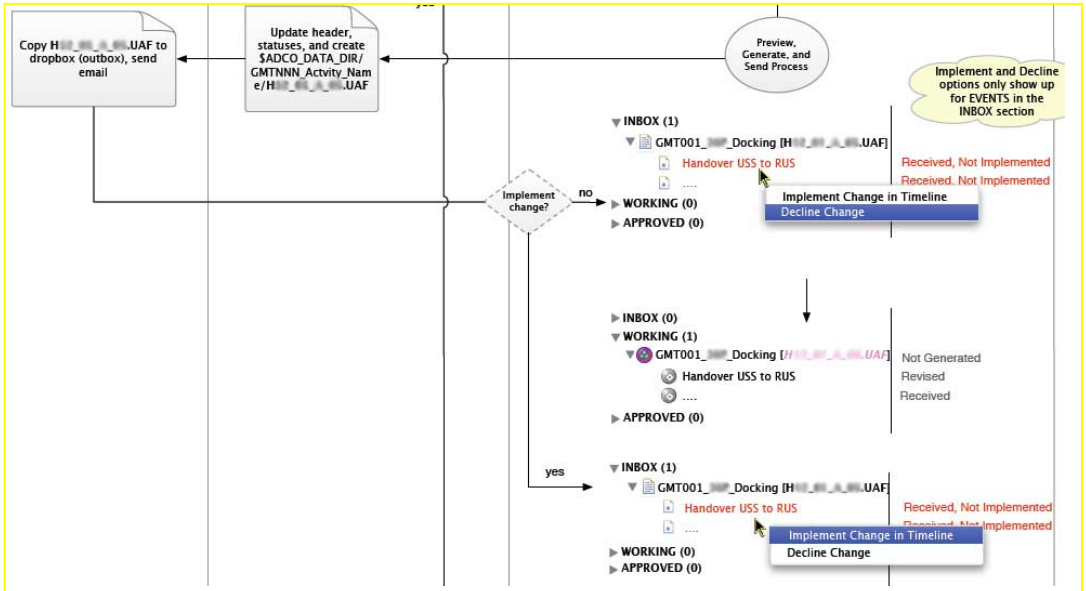


Figure 3. This is part of the design document for the file processing workflow (correlates with highlighted section in Figure 1 requirements document). The left columns indicate the state of the file system, and the right column illustrates the tool interface supporting user decision-making.

Major system-level designs were validated via cognitive walkthroughs using interactive low-fidelity prototypes (PowerPoint with images). After major frameworks were decided, implementation details (e.g. colors, formatting, button placement) were worked out between the product manager and ADCO project engineer. In making decisions, the ADCO project engineer liaised with the ADCO team, and the HCI group actively provided design and usability guidance and input.

2.4 Design, development and evaluation process.

The HCI group worked in 6-8 week development cycles. Prior to the start of the development cycle, the user research and design work already should have been negotiated and completed, allowing developers to accurately assess and estimate fully-developed designs. Several forms of representation were used to communicate the design details from the designer to developers, including workflow showing states of the file systems and corresponding user interface components being affected by the work (Figure 3) and schematics showing the decision structure in scenarios (Figure 2). After feature development, the product manager created testing suites based on the use cases to ensure the acceptability and functionality of the features.

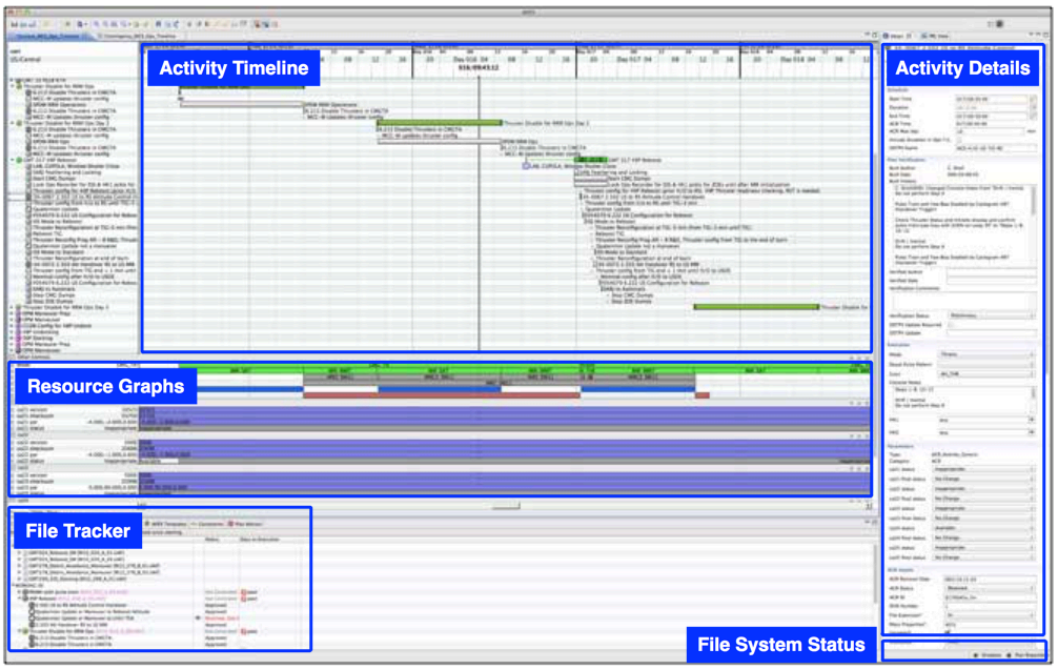


Figure 4. Screenshot of APEX with some major features delineated.

There were seven major releases in the two-year project. It was helpful to provide informal, intermediary releases to give users the opportunity to try out the features and point out problems or usability issues. For each release,

documentation of features was included and the ADCO test engineer was responsible for ensuring that the release met criteria. Acceptance criteria included meeting the items as listed in the requirements document as well as user satisfaction with the design and implementation of the features. Figure 4 shows a screen shot of the near fully-developed system.

3 Case 2: Scheduling Airlift Operations

Global airlift operations are complex and dynamic, requiring complex planning and coordination up to and through an operation. The Air Force customer for Case 2 wanted to improve support for scheduling operations and had previously worked with the provider group. The provider group was thus known to the customer, but not to the specific stakeholders in this particular airlift group. Stakeholder trust in and perception of value of the providers' work increased over the project, increasing the providers' access to operators. The provider group had worked together across many years and projects, and had well-established work practices.

3.1 Project context

This case study illustrates work-centered design (WCD). The phrase 'work-centered' is intended to highlight that the focus of the analysis is on the demands and broader context of the work (Eggleston, 2003). The project was conducted for an airlift organization that is responsible for scheduling and tracking airlift missions worldwide. The team developed a decision-support system to enable airlift operations center staff to understand and revise air-lift mission schedules 24 hours prior to and during mission execution (Roth, et al., 2006). This is one of several prototype support-systems this group developed for airlift mission planning and scheduling for the Air Force customer over the span of fifteen years (Roth, et al., in preparation). Typically prototype development projects are one to three years duration, with regular review from an integrated product development team made up of Air Force customer stakeholders. Each project ends with a formal user evaluation, to establish that users concur that the prototype provides effective support for their work, at which point the system is handed over to the Air Force customer for transition.

3.2 Team composition

A key aspect of the WCD approach was that a multi-disciplinary team participated in all phases of the design effort starting from the initial

knowledge capture sessions through to the final prototype evaluation. The core team included a cognitive engineer who is able to draw out sources of cognitive complexity and requirements for support through observation and interview of user groups, interface design specialists who are able to translate support requirements into effective graphic visualizations, and software engineers who are best able to grasp technological constraints and possibilities. The dialectic interplay among these multiple perspectives was critical to identifying opportunities for support that were simultaneously grounded in an understanding of what is needed and what is possible, and translating those support requirements into effective visualizations.

3.3 Requirements gathering process and work representation

A fundamental aspect of WCD is an analysis of the work 'context of use' to uncover the elements of work that require support. The process starts with knowledge capture methods such as ethnographic field observations and structured interview techniques (Bisantz and Roth, 2008) to uncover the characteristics of the work domain, the work requirements, the sources of complexity and cognitive and collaborative demands entailed. Formal methods are then used to represent the results of the analysis. These include work domain representations that capture the goals and constraints of the work and models of workflow within and across individuals and groups required to achieve work goals (see Bisantz and Roth, 2008 for further discussions and examples of work domain and workflow representations).

Figure 5 shows a portion of a work domain representation that was used to capture the factors that need to be considered in understanding and modifying a mission schedule. Work domain representations are used to provide a shared understanding and memory aid of the factors that need to be included in displays intended to aid support mission rescheduling. The representation is used both to communicate inside the design team as well as to communicate to users and stakeholders the current understanding of the design team of the factors that matter. While these representations are developed early in the project, they typically are not needed once the team becomes more versed in the domain.

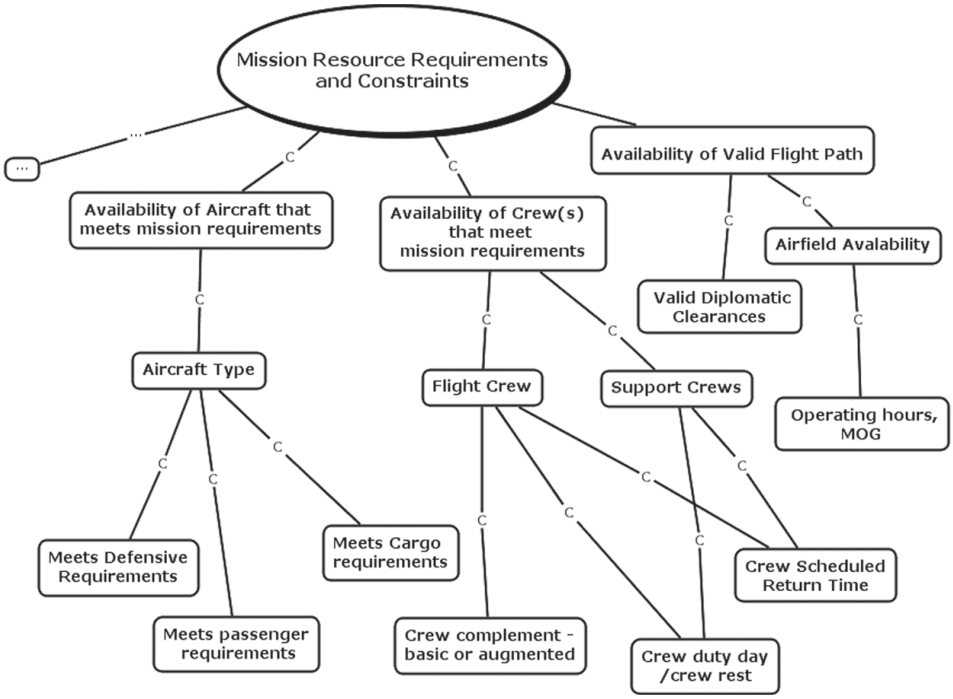


Figure 5. Selected portion of a work domain representation for flight scheduling (Taken from Evenson et al., 2008).

Throughout the development of the work-centered prototype, the design team produced informal products for communicating among themselves regarding the cognitive support requirements that needed to be met and how the evolving design would satisfy them. These included interview notes, notes describing implications for cognitive support requirements, and rapid prototypes of display support concepts.

Once they achieved consensus on the basic design features, they produced materials with which they could present stakeholders the design concepts, their prospective benefits, and scenarios illustrating how the concepts would be employed. A particularly useful intermediate product is a set of 'at a glance' cognitive work requirements that are often stated as questions that the user needs to be able to answer 'at a glance' by looking at the display. Figure 6 provides examples of 'at a glance' cognitive requirements. These cognitive requirements provide the design rational for more specific visual elements and forms that make up the display. These are referred to as 'work-centered display requirements'. Examples are provided in Figure 6. Wampler et al. (2006) provides a more detailed discussion of approaches to communicate cognitive requirements to support traceability from cognitive analysis to display requirements.

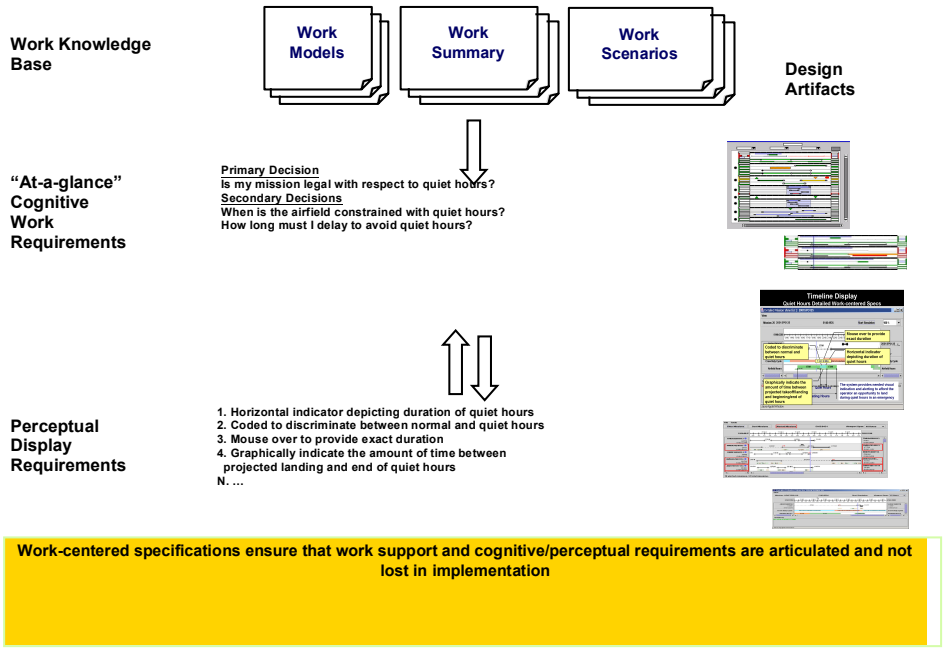


Figure 6. Example of an "at-a-glance" cognitive work requirement. (Taken from Wampler et al., 2006).

3.4 Design and development process

Based on the analysis, a novel timeline display was designed to provide visibility into the domain factors that impact mission viability (See Figure 7). It enables execution personnel to "see" the relationships between mission plan elements (e.g., planned mission take-off and landing times at different airfields) and resource constraints (e.g., airfield operating hours; durations of diplomatic clearances; crew rest requirements). This visualization allows operations center personnel to directly perceive the validity and robustness of a mission plan. Alerts are integrated into the visualization to highlight exceptions and guide problem-solving . In addition an active "what-if" simulation mode is provided to help assess alternative courses of action should mission problems arise (e.g., situations that create mission delays). The user can go into simulation mode and manually 'drag' a mission to a new position on the timeline to see what impact this change has. Alerts come up if the change creates any problems.

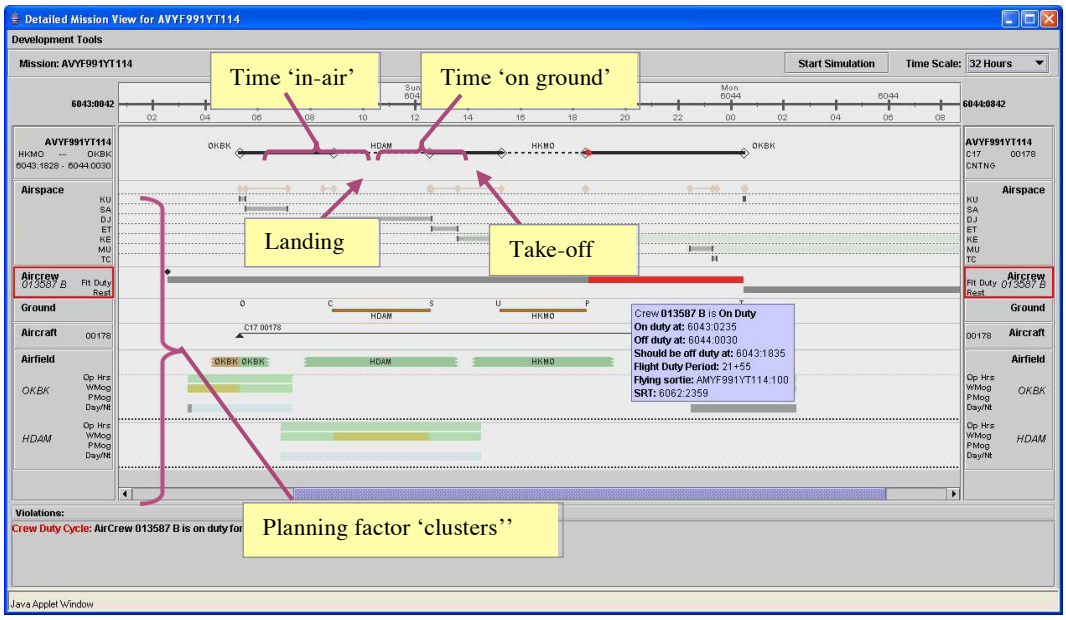


Figure 7. Screen shot of planning resources provided by the software (Taken from Wampler et al., 2006).

3.5 Evaluation and feedback

An evaluation study was conducted to compare user performance with the prototype timeline display vs. performance with the legacy system (Roth et al., 2006). Twelve experienced operations center personnel participated in the study. Test participants worked through different, but comparable, realistic work scenarios in the two test conditions. Results showed statistically significant improvement in terms of both faster solution times and reduced errors with the prototype timeline. Participants took approximately twice as long and made more than twice as many errors in the legacy condition than in the timeline condition. Self-reported ratings of situation awareness and workload were also significantly improved with the timeline vs. the legacy system. The timeline display is currently being transitioned by the airlift organization.

4 Commonalities and Differences

These two case studies were done by unrelated groups, unaware of each other's project, yet had some striking similarities. Concerning the projects, both supported planning and scheduling safety-critical movement, one of an extraordinarily complex vehicle, the ISS, the other of a vehicle fleet. In both cases plans developed over time, and re-planning was an important part of work. In both, plans were affected by technical and political factors, as well

as the dynamically adjusted goals to be accomplished. In both cases communication both within the planning group and to other groups affected by the plan were important. Both focused on enabling users to work effectively, safely, and efficiently.

Concerning the processes, both projects had important assets prior to project start. Both were done by or within established working groups, with developed work practices for identifying user needs in difficult work domains and for supporting those needs. Each provider group also had members quite familiar with the type of domain. Both provider groups had established relationships as providers to groups known by and related to the customer, thus increasing initial good will, the expectation of a valuable product, and management support. Both produced successful products. In the Airlift Case, the customer initiated a project to transition the design embodied in the prototype from the research/development group to be conducted by a commercial software development house. In the ISS Case, the implemented design has gone live in ISS Mission Control.

Within this strikingly similar framework, the two projects differed in many details of the processes used and the representations created by and supporting those processes. These are summarized in Tables 1 & 2.

Table 1: Comparison of Process

Processes & Roles	Case 1: ISS	Case 2: Airlift
Provider Roles & scope	1 product manager doing majority of user research and design; 4-5 part-time developers; involved project manager and larger group. Relatively distinct roles and hand-off processes. 2 yr project timeline.	1 cognitive engineer less that half time; 3-4 developers part to full time; additional affiliated SME part time. Highly interactive. Relatively informal and interactive processes. 3 yr project timeline.
Interaction with Users	Deep, on-going relations with one or two user-experts who liaised with user group as needed.	Emphasis on sampling multiple users.
Initial Assets-provider group	Established work practices in larger group	Shared collaborative history of all

	(HCI) but new product manager.	members.
Initial Assets-domain knowledge	Extensive prior experience designing and developing planning software. ADCO work largely unfamiliar.	All members familiar with domain-type and customer organization; previously developed a successfully fielded, work-centred prototype for customer.
Initial Assets-customer relations	Strong reputation known to stakeholders/users.	Strong relation with customer but not stakeholders.
Development structure	Structured Agile development cycles. Product manager completes and negotiates designs prior to hand-off to developers.	Informal Spiral Development. Integrated involvement of multi-disciplinary team throughout.

Table 2: Comparison of Representations

Representations	Case 1: ISS	Case 2: Airlift
Work domain/work needs	Largely informal notes; also indirectly in a) requirements document and b) design document for developers such as how workflow should be displayed.	Initially focal representations of work domain and workflow; less important when team had domain expertise as in this project
Requirements/ project commitments	Formal, stable Requirements specification. Defined scope of contract. Organized sequence of partial releases using agile development	Informal, primarily within group. Group built more formal requirements as a product to guide the developers to whom the project was

	cycle	transitioned.
Solutions, design concepts	Typically proposed multiple designs to discuss with user. Prototypes (interface behaviour in PowerPoint) for major features to confirm design.	Implemented components within the full design are shared to get feedback for one design "spiral"

5 Hypothesized Success Factors and Relevance to Agile Development

The assets at the start of each project were a big contributor to success, particularly development groups with strong work practices, oriented to the user (particularly Case 1) and the work (particularly Case 2). Both projects were moderate scale, but had different overall development structure. Given the similar initial assets, both development structures lead to a successful design.

Many projects will not start with these initial assets. Consider a provider group that has less domain knowledge, less established work practices, or less experience working as a team. Here, what practises or representations can increase the odds of a product that supports the user in carrying out the work, and thus provides a good user experience? We surmise that with developer groups that don't share these assets, the product will benefit a) from explicit and structured interaction within the provider group and between provider and stakeholders and b) with more structured use of explicit, shared, external representations.

The points where explicit representations provide most benefit may vary to some extent depending on where the informal communication channels are most fragile. However, for safety-critical and technically dense work explicit external representations are also very important as cognitive aids for organizing and transforming information, and serve both cognitive and communicative functions. How representation for oneself is related to representation for others, particularly those with less shared context is an interesting and little-explored topic (but see Jamieson, Miller, Ho, & Vicente, 2007). We comment on representations from the perspectives of content topic and of representation use across contexts spanning the needs-analysis through evaluation cycle.

5.1 Types of Representation Content

We group representation types into three very general, higher-level topics and mention examples of each type used in the case studies: representations of the work, representations of the design, and representations of how the design supports the work.

1) Explicit, direct representations of the work were particularly important for both groups in projects prior to the current case studies, when each group was unfamiliar with relevant planning domains. The Case 1 group used Contextual Inquiry and the Case 2 group used Cognitive Work Analysis. In the Case 1 study, the Requirements was an important, stable document and primarily expressed work functions that had to be supported by APEX. Thus, this form of Requirements did much to represent the work domain.

2) Representations of design are particular important if the person designing is 'handing off' the design to programmers, as happened in Case 1. The multi-representation "design document" included specifications showing a work function and how it was to be supported.

3) The relations between design choice(s) and the work function(s) supported by them is valuable information. The relations are particularly important to represent explicitly if a group unrelated to the initial analyst/ designer group will be doing the programming. The Case 2 group wrote *requirements* as *output* documents, the set of 'at a glance' representations. The 'at a glance' requirements link cognitive work to a constrained set of allowable design options, that is, those that would preserve the design rationale of the prototype that had been developed by the group. This was critical to protect the cognitive engineering for transitioning the prototype to a production software house.

5.2 Construction and Use of Representations across Contexts

It is very challenging to embody the knowledge resulting from good analysis and design into representations useful for a variety of users and circumstances. For example, often information can best be understood in small components (scenarios, work functions, displays), yet ensuring the coherence of an overall design and its support across the range of intended work functions requires an overall perspective. Representations which facilitate transition from focus on a component to focus on larger or overall structure may be helpful (Billman, Archedeacon, et al., 2015; Pritchett, Kim, & Feigh, 2014) and also visualizations which provide global-local shifts. Ideally, the same content can be provided to guide design and used in evaluation, such that design is guided

by the evaluation criteria. Representations most suited for formal methods, as in verification, may not be suited for human inspection, as primarily needed for validation. In addition, the up-front costs of building representations must be balanced against later benefits, often for other people, of having the information explicit and accessible.

Improving the usefulness and hence user experience in technical work domains hinges in large part on understanding the work to be done. The two cases presented here are examples of projects that succeeded, at least in the relative sense of providing much better solutions than legacy software. Both cases developed and depended on effective representations of work needs and of designs to address those needs. Understanding of and development of alternative representations and conditions will be important for reliably producing software for technical work domains that is effective, and thus, produces good user experience.

Acknowledgements

Authors are ordered alphabetically. The authors wish to acknowledge the contributions of JSC ADCO and the Ensemble team, which includes members from the Human-Computer Interaction (HCI) group within the Human Systems Integration Division (Code TH) and the Planning and Scheduling group (Code TI) that contributed to the design, implementation, testing, support, and delivery of the software tool. Part of this work was supported by funding from NASA Human Research Program: Space Human Factors Engineering (466199.02.01). We would also like to acknowledge the interdisciplinary team that contributed to the analysis and design of the airlift scheduling systems.

References

[1] Aghevli, A., Bencomo, A., & McCurdy, M. (2011). Scheduling and Planning Interface for Exploration (SPIFe). Presented at the International Conference on Automated Planning and Scheduling (ICAPS), Freiburg, Germany.

[2] Billman, D., Archedeacon, J., Christopher, B., Deshmukh, R., Ebbs, K., Fan, C., … Stewart, M. (2015). *Needs Assessment and Work Allocation Tools for Mission Operations and Procedures*. NASA Ames Research Center.

[3] Billman, D., Arsintescu, L., Feary, M., Lee, J., & Tiwary, R. (2015). Needs Analysis and Technology Evaluation: Evaluation Case Study of Alternative Software for Controller Planning Work-Part 2. *Journal of Cognitive Engineering and Decision Making*, *9*(2), 186–207. http://doi.org/10.1177/1555343414567775

[4] Billman, D., Feary, M., Schreckenghost, D., & Sherry, L. (2015). Needs Analysis and Technology Alignment Method: A Case Study of Planning Work in an International Space Station Controller Group-Part 1. *Journal of Cognitive Engineering and Decision Making*, *9*(2), 169–185. http://doi.org/10.1177/1555343414567774

[5] Bisantz, A., & Roth, E. M. (2008). Analysis of Cognitive Work. In D. A. Boehm-Davis (Ed.), *Reviews of Human Factors and Ergonomics* (Vol. 3). Santa Monica CA: Human Factors and Ergonomics Society.

[6] Eggleston, R. G. (2003). Work-Centered Design: A Cognitive Engineering Approach to System Design. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *47*(3), 263–267. http://doi.org/10.1177/154193120304700303

[7] Evenson, S., Muller, M. and Roth, E. M. (2008). Capturing the context of use to inform system design. *Journal of Cognitive Engineering and Decision Making*, 2 (3), 181-203.

[8] Gross, T., Gulliksen, J., Oestreicher, L., Palanque, P., Oliveira Patres, R. and Winckler, M., eds. Proceedings of the 12th IFIP TC.13 International Conference on Human-Computer Interaction - INTERACT 2009. Springer-Verlag, Heidelberg, 2009.

[9] Holtzblatt, K., Wendell, J. B., & Wood, S. (2005). *Rapid contextual design : a how-to guide to key techniques for user-centered design*. San Francisco: Elsevier/Morgan Kaufmann. Retrieved from http://www.loc.gov/catdir/toc/els051/2004061522.html http://www.loc.gov/catdir/description/els051/2004061522.html

[10] Jamieson, G. A., Miller, C. A., Ho, W. H., & Vicente, K. J. (2007). Integrating Task- and Work Domain-Based Work Analyses in Ecological Interface Design: A Process Control Case Study. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, *37*(6), 887–905. http://doi.org/10.1109/TSMCA.2007.904736

[11] Pritchett, A. R., Kim, S. Y., & Feigh, K. M. (2014). Measuring Human-Automation Function Allocation. *Journal of Cognitive Engineering and Decision Making*, *8*(1), 52–77. http://doi.org/10.1177/1555343413490166

[12] Roth, E. M., DePass, B., Scott, R., Truxler, R., Smith, S. Wampler, J. (in preparation). Designing collaborative planning systems: Putting Joint Cognitive Systems Principles to Practice. In P. Smith and R. R. Hoffman (Eds). Cognitive Systems Engineering: *The Future for a Changing World*. Ashgate Publishing Limited.

[13] Roth, E. M., Stilson, M., Scott, R., Whitaker, R., Kazmierczak, T., Thomas-Meyers, G. and Wampler, J. (2006). Work-centered design and evaluation of a C2 Visualization Aid. *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting*. (pp. 255- 259). Santa Monica, CA: Human Factors and Ergonomics Society.

[14] Wampler, J., Roth, E., Whitaker, R., Kendall, C., Stilson, M., Thomas-Meyers, G., and Scott, R. (2006). Using work-centered specifications to integrate cognitive requirements into software development. *Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting*. (pp.240- 244). Santa Monica, CA: Human Factors and Ergonomics Society.

A Multi-Method Approach to UX Requirements: Adapting to Agile & Lean Development

Gregorio Convertino
Informatica Corporation
2100 Seaport Blvd, Redwood City, CA 94063, USA
gconvetino@informatica.com

Abstract. Agile and lean development projects are becoming a common practice among software development teams in industry. This new context poses adaptation challenges for traditional UX requirements methods. This paper summarizes a case study on coordinating multiple methods for requirements specification and validation during agile and lean development. In particular, it identifies four specific challenges, presents a commented list of the UX methods used with their benefits, and finally presents a few lessons learned on ways to orchestrate multiple UX methods as part of a program of research that runs along the development lifecycle. The 20-month development project here presented as context for such methods delivered Informatica Rev, an innovative software tool for data analytics by business users.

1 Problem

The *context* of this case study is a 20-month development project at Informatica Corporation that delivered Rev [9] (Figure 1). Informatica is a global software company with about five thousands employees. It develops innovative enterprise software for big data integration and analytics.

Rev is an innovative software tool to enable analysts in businesses to perform self-service data preparation (e.g., data analysts Sales and Marketing departments). The development of this tool is an example of an agile and lean process in an enterprise that needs heavy requirements specification and validation efforts. Two key factors motivated this need as the project started: new users and a new market. Data analysts are a new class of users, or unfamiliar personas, to Informatica, the company that built Rev. Moreover, data preparation applications are a new emerging market with no clear tools of reference (e.g., [6]).

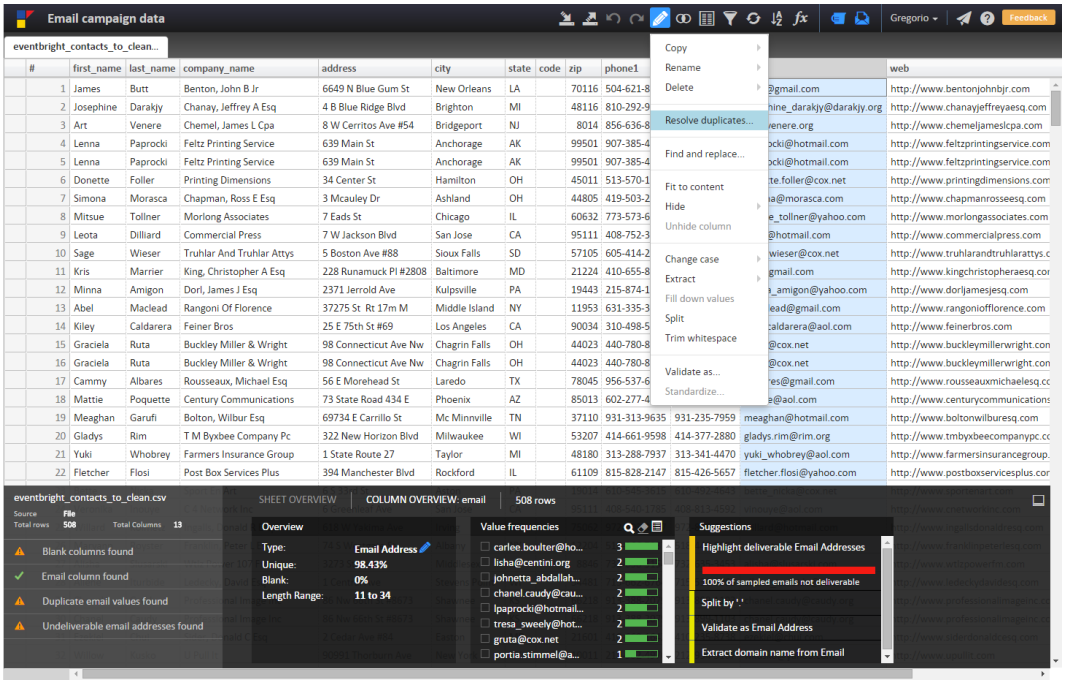


Figure 1. Rev Product. The reader is invited to try out the software for free via the URL in [9].

The *audience* of this case study are primarily professionals in user research, professionals in software design who need good requirements, and academicians teaching courses relevant to user research or software design roles. The practices presented are particularly relevant to large organizations that have integrated the roles of UX designer and UX researcher (or user researcher) as part of their agile development teams.

Consistent with the themes of the workshop, the *problem* tackled in this position paper stems from the practical challenges of applying traditional UX requirements methods to agile and lean product development. We outline below four challenges that user research professionals face today in this context. The first three are about effective UX method selection and orchestration; the fourth is about balancing different types of requirements.

1) *Some traditional UX methods are not compatible with the fast and iterative nature of agile and lean development processes* (e.g., field studies, ethnography, and large surveys). Thus adaptations of these methods need be developed to work with small-N studies and incremental research programs (e.g., [8]). In this case study, studies were run taking into account a development cycle with iterations or sprints lasting six to eight weeks, which suggests that the requirements studies should be planned, run, and delivered within one or two sprints periods. The use of a sprint periods of six to eight weeks, while

relatively long for agile development standards, it was required given the novelty and complexity of the Rev product.

2) *Traditional UX requirement methods are often used prescriptively as fixed-stage, fixed-purpose instruments,* rather than as mere "knowledge extraction instruments" that can inform design at *different stages and with different purposes*. The latter approach gives the UX researcher more flexibility to plan the investigations of requirements. For example, online surveys can be useful to supplement interviews early in the process, characterize new user classes later in the process, or validate designs after they are implemented. This is not intended to deny that methods can be more appropriate at given stages or for given purposes but it points to challenge for the UX researcher to distinguish the study method from the stage and purpose of the study. Making this distinction gives more flexibility for doing method selection or adaptation during the development cycle, which leads us to the next challenge.
3) There are various effective manuals that provide useful inventories of methods for UX requirements (e.g., [1]; [3]; [7]). However such methods are often proposed or mastered as atomic solutions. *The implicit assumption is that method selection and planning is done at the level of a 'single study' rather than as a 'program of user research',* which covers the full development lifecycle from early requirements and concepts to released product. Two observations from daily research practice falsify this assumption. First, requirements studies run within the same project are inevitably connected and thus decisions about methods of individual studies are impacted. For example, to address subsequent requirements questions, the UX researcher can run studies that apply the same method or *different but compatible* methods if these are more convenient (e.g., a survey with users can be an efficient way to extend the requirement specification work done in an earlier study that used interviews and observations with the same users). Second, the fast pace of lean development requires the UX researcher to plan incrementally and break the research efforts in smaller chunks, which makes the need to plan at the level of the research program even more evident: i.e., striving to *connect and coordinate concurrent methods* (e.g., two complementary methods) *and subsequent methods* (e.g., an incremental survey across iterations or sprints). The reader can also find a few additional insights on ways to connect qualitative and quantitative methods in [4].
4) When planning work on UX requirements and later when deriving implications for design, the UX researcher and UX designer, working

together, *need to accommodate non-user requirements that will inevitably steer the design and development* (see "Business Requirements" in [1]). Examples of non-user requirements in industry include: the company's product strategy (e.g., incorporating new features or capabilities in order to integrate with existing products by the same company), "appealing" features that the market requests via customers (i.e., when these are decision makers but not users), managers, or market analysts with influence on the product leadership (e.g., features that competitors are offering), and non-core features needed to increase adoption via partner companies (e.g., features to integrate with existing tools already used by analysts). The challenge for the UX researcher and the designer is to strike a good balance between user and non-user requirements, while protecting the consistency of the overall design.

To address the first three challenges above, the following sections present the case study in steps. The next section provides a chronology of the development project as context. Then, the following section presents a commented list of the methods used, which address the first two challenges. Finally, the paper presents four lessons learned on ways to orchestrate UX methods as part of a program of research that runs along the development lifecycle.

2 Context: Chronology of an Agile and Lean Project

Figure 2 provides an overview of the multiple user research methods used over 20 months, from December 2013 to June 2015, in support of the agile and lean software development process. The process led to Informatica Rev, a new software tool for business analysts (Figure 1). Besides the methods used to evaluate or refine designs (prototypes or live tool versions), multiple methods (e.g., workshops, user studies, and surveys) were also used to progressively scope, specify, and validate user requirements.

The requirements specification and validation process was launched at the end of November 2013. The team that worked on requirements included various professional roles: product manager, UX designer, user (or UX) researcher, senior user interface engineer, and a technical architect. In about three weeks, under the guidance of the product manager this small core team outlined the vision of the new software product and the hypothetical personas. In particular, the team consolidated the early requirement definitions during a multi-day workshop. The set of features envisioned for the new product was specified through a user story map [5]. In addition, an early description of two

hypothetical personas ("Line-of-business data analyst" and "Data hero") was outlined based on data collected by the product manager over the prior weeks (see top left in Figure 2).

The story map and the hypothetical personas became the starting point for the UX researcher and UX designer to run a field study with seven target users. This study specified and validated user requirements. In particular, it provides details about the target users, identifying the core persona (Appendix, Persona: Ops Data Analyst), specified their data preparation tasks (i.e., analysis projects and specific data manipulations) and key unfulfilled needs during data preparation. To adapt to the fast pace of the agile team, the small field study was planned, run, and its findings were reported within about six weeks, by mid-January 2014 (see challenges in the section above) (Study 1 in Figure 2).

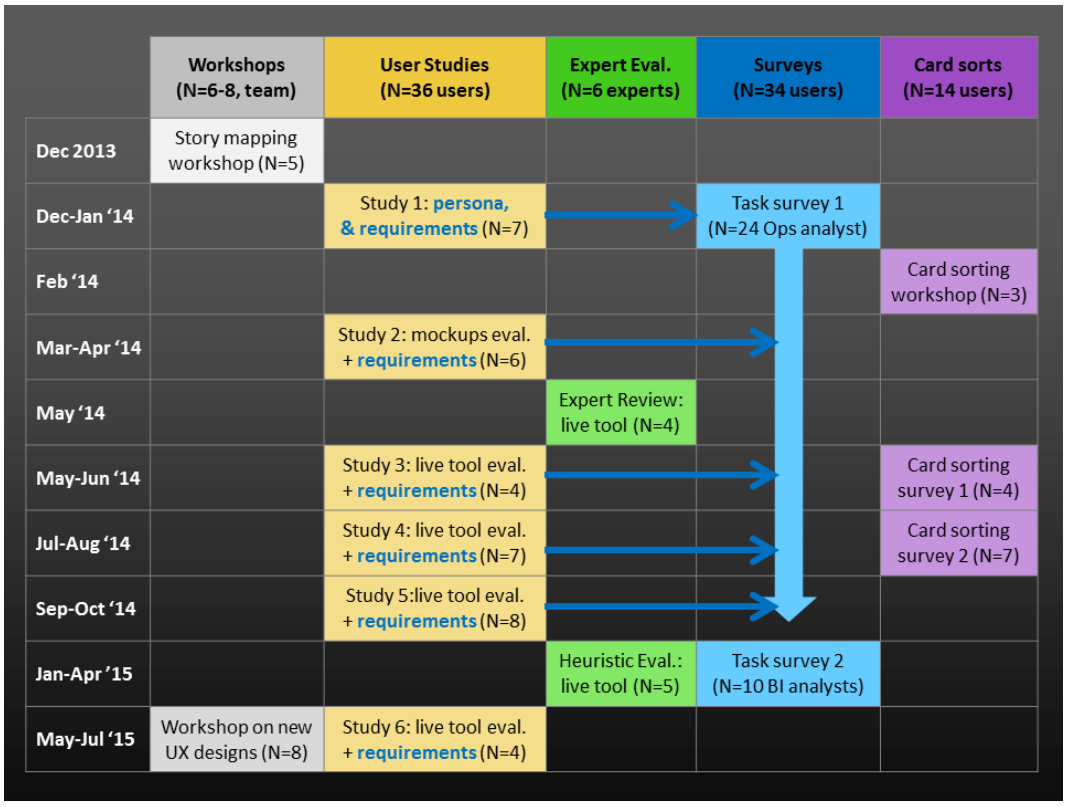


Figure 2. Timeline by UX method. The figure lists multiple user research methods used over 20 months, from December 2013 to June 2015, in support of an agile and lean software development process that led to Informatica Rev (see Figure 1).

Over the following two months, this requirement study helped the team focus its priorities. First, it gave the team a detailed description of the primary persona (e.g., with specific tools used, skills, and differentiated from the secondary persona of the data scientist). Second, it provided a short list of common data analysis questions, data manipulations, and data preparation

issues that the study participants faced (e.g., resolving duplicates, joining data tables), which helped to prioritize the feature to be prototyped. Third, these findings helped the team define two realistic and relevant scenarios for the envisioned tool. These scenarios became the context for a set of prototypes by the UX designer: storyboards with wireframes (e.g., how to enables duplicate resolution in Rev).

As the design converged on a first stable prototype, in the form of a storyboard, a first design validation study was planned and conducted with a small sample of six analysts, between March and April 2014 (Figure 2: Study 2).

Then, during the following year from May 2014 to April 2015 six subsequent design validations studies were run every time a new version of the live tool was released in the development process. These included four small-N studies with users (Figure 2: Studies 3-6) and two expert evaluations with UX experts (Figure 2: Expert Review and (adaptation of) Heuristic Evaluation). Concurrently, a small workshop with three end users and then two follow-up card sorting surveys were run to improve the subsequent versions of toolbar and menus.

While these studies focused on validating designs, from January to October 2014 the requirement specification related to the core persona (see Appendix, Persona: Ops Data Analyst) continued through a new method arrangement involving two complementary methods: a 20-minute interview about user tasks and needs preceded the design validation in each user study (Studies 1-5 in Figure 2), and an online survey (Task survey 1 in Figure 2).

In January 2015, a new key persona was discovered as relevant to the new tool. Thus, from January to April 2015, a second survey collected requirements about this new persona: Business Intelligence (BI) data analyst (Task survey 2 in Figure 2).

Finally, between May and July 2015, after a new user study on new features of the live tool, the team ran a second workshop to update the shared knowledge about personas and requirements (i.e., core personas with their key use cases and needs) and thus review what areas of the tool design required rethinking.

3 UX Methods Used: a Commented List

The reported uses of UX methods such as workshop, user studies, and surveys address primarily the first two challenges described in the Problem section. This section includes a list of methods that proved effective in the agile and

lean context and shows how some methods can serve different goals at different stages of the development cycle (e.g., user study, survey).

3.1 Workshop: Story Mapping and Persona Drafting

The first workshop involved four core members: product manager, UX designer, UX researcher, and a software development leader. The product manager outlined the product vision and then the team defining the user story map as initial characterization of the space of capabilities envisioned for the new tool (see Appendix, Persona: Ops Data Analyst). The team generated a user story map following the method proposed by Patton (2005). The workshop and subsequent conversations allowed the UX researcher to outline two hypothetical personas and user needs to be validated during user design. The personas were drafted using the method proposed by Gothelf and Seiden [2].

After about a year and half of design, user research, and development, the second workshop was run primarily for the team to update the shared knowledge on requirements (i.e., the core personas with their key use cases and needs) and identify what areas of the tool design require refinement or rethinking.

Workshops have useful advantages as UX requirement method:

- It is fast, typically can be completed a few days.
- The team can share understanding and build consensus quickly.
- The team can discuss personas in depth.
- The team can quickly unpack use cases and discuss their priorities.
- The team can understand competitive products and their UX.
- Facilitate rapid, iterative concept generation.

In brief, workshop can serve as a springboard to plan the design and research deliverables for the following iterations or sprints.

3.2 User Study for Persona and Requirements Specification

This first user study, run with seven data analysts, involved semi-structured in-person interviews and observations in the participant's office. The data analysts were recruited based on screening criteria agreed beforehand with the Product Manager and the UX Designer (Study 1 in Figure 2). The first outcome of this study was the validation of the two hypothetical personas: we used the persona format proposed by Gothelf and Seiden [2], as it lends itself well for lean UX. The second outcome was the specification and validation of requirements in the form of common analysis questions, tools used, data manipulation operations, and data preparation issues.

3.3 User Studies for Design Validation plus Requirements Specification

This second class of user studies involved semi-structured interviews, design walkthroughs, and recording of the user's screen during the interview session. These were always aimed at two goals. The primary goal was to evaluate the design. The secondary goal was to continue specifying requirements and characterizing users with their data preparation issues (Studies 2-6 in Figure 2).

3.4 Expert Reviews

In this case study, expert reviews proved to be a useful UX method to rapidly identify key issues in the design and collect principled recommendations on how to address these issues. As soon as the first live version of the tool was released, the UX researcher run a first expert review, which involved four experts. It took four weeks and provided quick feedback. Then, about a year later, a more systematic version of expert review was used, an adaptation of the heuristic evaluation method, which involved pre-specified heuristics and tasks (Appendix: Figure 7).

Expert reviews are a particularly useful in an agile and lean development when interleaved with user studies because the experts offer a type of feedback that complements the feedback from the target users. The feedback from the users who evaluate the tool ensures that the design is useful and effective: i.e., it addresses real requirements of real users in real task settings. The feedback from experts ensures that the design is simple and harmonious: i.e., it follows known design principles (e.g., easy navigation, clear visual design) and UX patterns expected for the category of the tool evaluated (i.e., tools for data analysis) (Figure 2: Expert Review and (adaptation of) Heuristic Evaluation).

3.5 Survey Studies

The two survey studies used the online survey method to address different requirements specification goals, at different stages of the development lifecycle. Each survey required about 20-25 minutes to be completed. It included Likert scales (i.e., frequency and importance of specific data preparation tasks) and a few open-ended questions (e.g., top three data analysis questions answered) (see sample questions in Appendix: Figure 8).

The first survey run in parallel with and supplemented the requirements specification process during the user studies, which contained brief interviews on tasks and needs (Studies 2-5 in Figure 2). This survey allowed the

researcher to systematically collect information on the profile of the users who had participated in user study (Studies 2-5 in Figure 2) and also quantitative ratings about the importance of specific tasks and needs (Task survey 1 in Figure 2).

Differently from the first survey, the second survey was conducted independently from the user studies. It addressed a specification needed later in the development process: a new persona was discovered as relevant to the new tool. This second survey used questions analogous to the first survey to allow comparing requirements (e.g., the most important data preparation tasks) between the new persona of BI analyst and the old core persona of Ops Data Analyst (Task survey 2 in Figure 2; Appendix, Persona: Ops Data Analyst).

3.6 Card sorts

This method was used to inform the design of the toolbar and the menus. A first workshop with a group of three end users who met in person was followed by two subsequent card sorting surveys each providing new insights on how to improve the toolbar and the menus (Card sorting workshop and Card sorting surveys in Figure 2).

4 Lessons on Orchestrating UX Methods

The lessons proposed below offer a few examples of helpful orchestrations of multiple UX methods during a lean and agile development project. It is worth noting that the lessons are the results of reflections by the author from this 20-months case study. They were shared and distilled during various discussions with different collaborators in the UX team at Informatica, one internal and two external to the Rev team. These lessons should be taken as work in progress.

While based on this case study only, the lessons represent a first set of solutions to the third (and hardest) challenge described in the Problem section. That is, they suggest ways for the researcher to connect UX requirements methods in a "pipeline" that is planned at the level of program of research (Figure 3).

4.1 Lesson 1: Transition from Team Workshop on Requirements to a Small-N Field Study on Requirements

A first lesson learned was that it was helpful to rapidly transition from early requirements drafting by the team during the initial workshop (the story map

and the draft of two hypothetical personas) to a small-N field study that validated key parts of these requirements. See section above for more details on the methods. See point 1 in Figure 3.

4.2 Lesson 2: Transition from Field Study to Remote User Study plus Surveys

Running field studies with in-person observations is often too costly or simply not feasible when the users are geographically distributed. The second lesson is that remote user studies (using interviews and screen recording) are an efficient way to continue the requirements specification effort started with the initial small-N field study. During an agile and lean development process, as the researcher sets up small-N user studies for validating subsequent design, it is useful and cost-effective to allocate about 15-20 minutes to a brief interview on tasks and issues, before the user starts the design validation session (Studies 2-6 in Figure 2). See point 2 in Figure 3.

4.3 Lesson 3: Combine Small-N User Studies with a Survey

A key limitation of user studies during agile and lean development projects is the small number of participants, or small N problem. A survey can address this limitation: i.e., systematically validate requirements with enough users to ensure external validity.

During the case study this result was achieved by combining subsequent small-N user studies with a concurrent online survey. The design of a survey became possible based on the initial field study (Study 1 in Figure 2) plus with prior research on the same class of users (i.e., data analysts), which had provided the researcher with enough variety of data preparation tasks and issues to construct an online survey. The data collected using the survey were then interpreted together with the qualitative data collected in the short interviews (Studies 2-5 in Figure 2, blue arrows, or point 3 in Figure 3), see lesson 2 above.

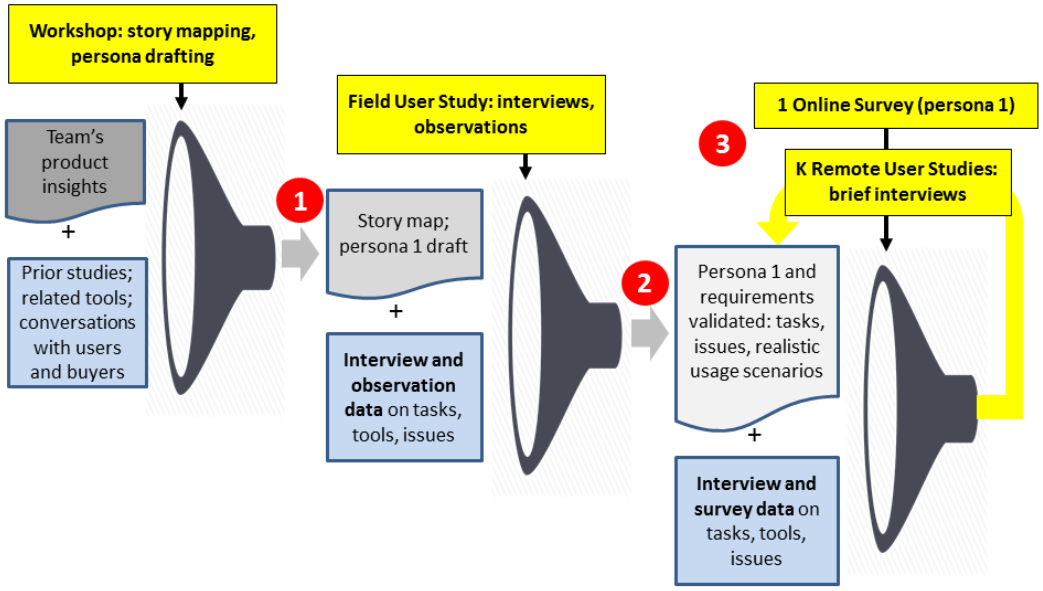


Figure 3. The pipeline model. The figure shows three UX methods to specify and validate requirements (workshop, field user study, and multiple (K) remote user studies combined with a single online survey) connected as part of a program of research. Each method uses multiple inputs to generate a more refined specification of the requirements (see grey boxes). *The three lessons described above correspond to the red labels 1, 2, and 3 respectively*. See also steps in Figure 2 from December 2013 to October 2014.

Figure 2 summarizes how three subsequent UX methods were connected to specify and validate requirements:

- Workshop (story mapping and persona 1 drafting)
- Field user study (interviews and observations)
- Multiple remote user studies (brief interview) combined with a single online survey.

The specification of requirements for the second persona, or persona 2, occurred after step 3 and is not represented in Figure 2.

4.4 Lesson 4: Use Follow-up Surveys to Profile New Personas

Later in the project, the use of a second survey proved useful for a different purpose: profiling a new persona (Task survey 2 in Figure 2). It allowed the team to systematically compare requirements (e.g., the most important data preparation tasks) between samples of users that represent two different personas, one known and one unknown. This was possible because the second

survey used questions that were analogous the first survey, with a few minor additions.

5 Summary and Conclusion

This paper presented a case study on coordinating multiple methods for UX requirements specification and validation during 20-month agile and lean development of an innovative software tool for data analysts. The paper makes three contributions.

First, it identifies *four challenges for traditional UX requirements methods* in agile and lean development

Second, it presents the *different UX methods used with the observed benefits*:

- Workshop: Story Mapping and Persona Drafting
- User Study for Persona and Requirements Specification
- User Studies for Design Validation plus Requirements Specification
- Expert Reviews
- Survey Studies
- Card sorts

Third, it proposes *four lessons on ways to orchestrate these multiple UX methods as part of a program of research* that runs along the development lifecycle:

- Lesson 1: Transition from Team Workshop on Requirements to a Small-N Field Study on Requirements
- Lesson 2: Transition from Field Study to Remote User Study plus Surveys
- Lesson 3: Combine Small-N User Studies with a Survey
- Lesson 4: Use Follow-up Surveys to Profile New Personas

Finally, two caveats for the reader. First, the lessons reported should be considered work in progress. They are the results of direct experience and post-hoc discussions with a few collaborators during the 20-month case study.

Second, there are fundamental limitations of Agile itself that were not discussed in this paper. Primarily, while Agile focuses on how to efficiently manage software development, the work on UX requirements, designs, and validation studies tends to be left off the table. But this goes beyond the scope of this paper.

6 Appendix

6.1 Story Map (sample)

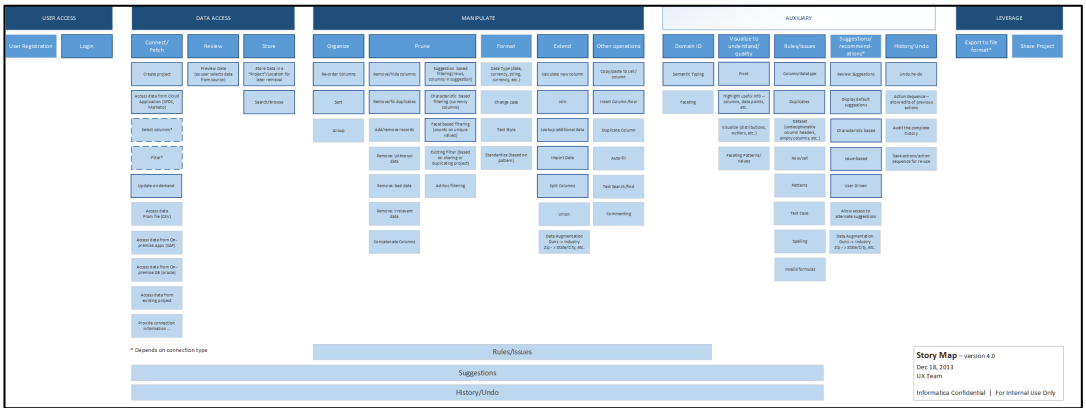


Figure 4. User story map. This is one of the user story map versions generated by the team during the initial workshop (December 2013).

6.2 Study 1. Persona 1: Ops Data Analyst – Mary

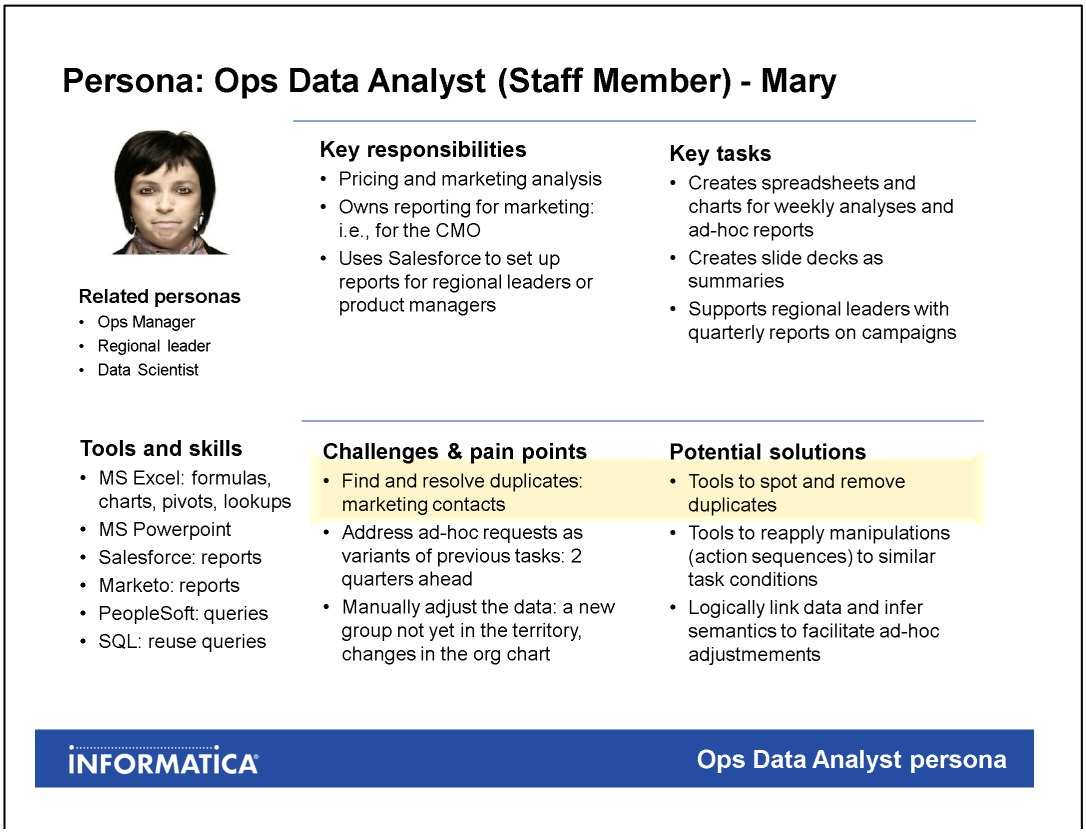


Figure 5. Persona: Operations (Ops) Data Analyst (e.g., Staff Member in Marketing Operations or Sales Operations). This is a version of the primary persona for the product delivered to the team as result of Study 1 (January 2014). It follows the persona description format proposed by Gothelf and Seiden [2].

6.3 Study 1. Requirements report (sample of findings)

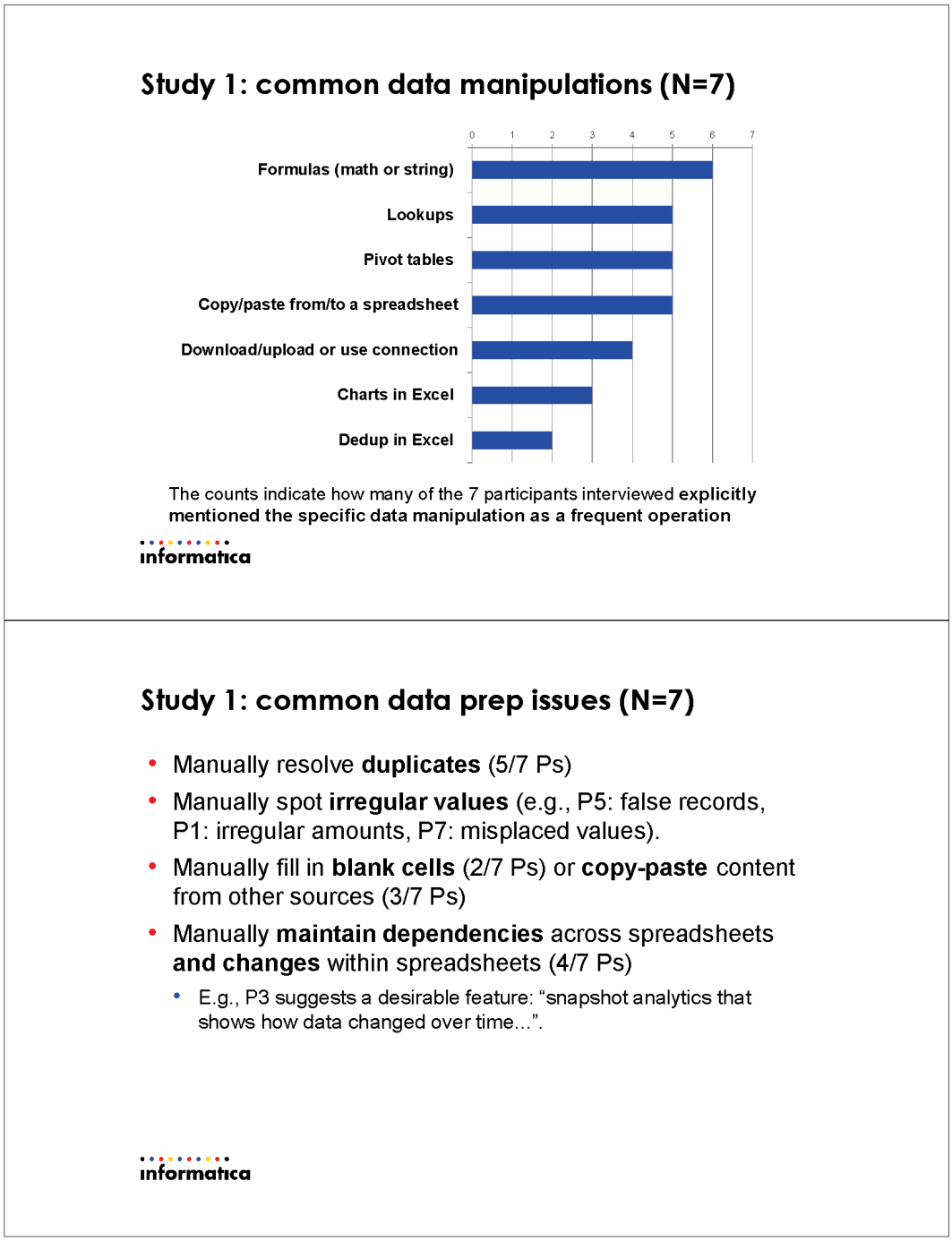


Figure 6. Study 1 requirements report: sample of findings presented to the team (January 2014). The top slide summarizes the common data preparation operations explicitly mentioned by the participants interviewed: i.e., operations to be supported. The bottom summarizes common data preparation needs that are poorly addressed by their current tools.

6.4 Heuristic Evaluation: an Adaptation. Heuristics and Tasks

		Heuristics
Generic	H1	Easy navigation & interaction (consistent and with feedback)
	H2	Clear visual design (controls, terminology, visualizations, layout)
	H3	Effective task flow or organization of operations
Prep-specific	H4	The UX patterns are consistent with common user expectations
	H5	The current prep task context is easy to understand
	H6	The next data prep steps possible are easy to predict
	H7	The tool supports data prep efficiency though automation and reuse
	H8	The tool includes key functionality expected in a data prep application

	Tasks
T1	Load, Understand, Improve quality
T2	Integrate data sets, Consolidate, Answer
T3	Add column, Aggregate, and Share via email
T4	Refresh aggregation based on new data and export the sheet

Figure 7. The second expert review used an adaptation of the Heuristic Evaluation method. It included the use of eight heuristics assessed by the experts in the context of each of four tasks. The heuristics and the tasks are listed above.

6.5 Survey Studies. Online Survey (sample of questions)

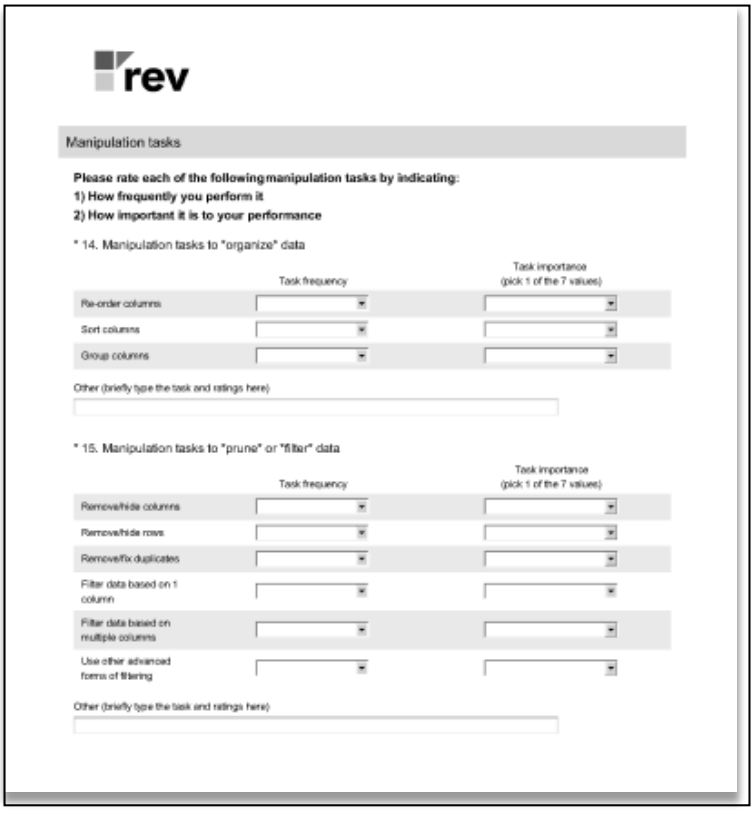


Figure 8. Online Survey. This method was used in two survey studies to collect requirements from each of two distinct personas for the Rev product.

Acknowledgements

I thank my UX colleagues Mohini Wettasinghe, Mark Detweiler, and Liam Friedland for the feedback on the paper and our multiple conversations, which helped me distil the lessons reported in this paper. I also thank the Rev team at Informatica Corporation since all the studies presented were planned and conducted in close collaboration with the team.

References

[1] Courage, C. and Baxter, K. 2004. *Understanding Your Users: A Practical Guide to User Requirements Methods, Tools, and Techniques*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

[2] Gothelf, J. and Seiden, J. *Lean UX: Applying Lean Principles to Improve User Experience*. O'Reilly. 2013.

[3] Hartson, R. and Pyla, P. *The UX Book: Process and Guidelines for Ensuring a Quality User Experience*. Morgan Kaufman, 2012.

[4] Klein, L. *Using Qualitative & Quantitative Research to Build Better Products*, O'Reilly 2015.

[5] Patton J. It's All in How You Sliece: Design your project in working layers to avoid half-baked incremental releases. *Manage People & Projects* www.agileproductdesign.com/presentations/user_story_mapping/

[6] Randall L. and Beyer M. A., 2014. Gartner Data Preparation Is Not an Afterthought. *Report G00270217. Gartner, Inc.,* November 7, 2014

[7] Rosson, M.B. and Carroll, J.M 2001. *Usability Engineering: Scenario-Based Development of Human-Computer Interaction*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

[8] Shapiro, D. 1994. The limits of ethnography: combining social sciences for CSCW. In *Proceedings of the 1994 ACM conference on Computer supported cooperative work* (CSCW '94). ACM, New York, NY, USA, 417-428.

[9] Informatica Rev, 2015. https://rev.informatica.com. To try out the software please register using this URL: https://rev.informatica.com/?code=code54

Requirements Elicitation for New Video Game Development Tools: A Case Study

Christos Fidas[1,2], Nikolaos Avouris[1], Ivan Orvieto[3]
[1]Human-Computer Interaction Group, [2]Dept. of Cultural Heritage and New Technologies, University of Patras, Greece, [3]Testaluna, Italy
fidas@upatras.gr,avouris@upatras.gr,orvieto@testaluna.it

Abstract. This paper presents a case study involving requirements elicitation for new tools in video games development. Eight video game developer companies from three different countries and a variety of stakeholders (N=17) participated in a user study that was based on a tailor-made requirements elicitation framework. During the process, interesting issues emerged related to the applied method but as well concerning the presentation of innovative tools, as disruptive technology, the different stakeholders' points of view and roles in the process, the role of technology providers and the organizational challenges towards this new game development pipeline. This case study provides interesting insights in applying a user-centered approach for requirements elicitation in the video game application domain and discusses lessons learned which can be of value for UX researchers and practitioners in video game research.

1 Introduction

In requirements engineering, requirements elicitation is the practice of collecting the requirements of a system from users, customers and other stakeholders. The practice is also sometimes referred to as "requirement gathering" [1]. During the last decade there has been a growing interest in the adaptation and customization of requirements elicitation methods and techniques to the unique characteristics to each individual application domain (e.g. mobile learning, serious games, banking systems) [2-4]. The attempt to bootstrap requirements elicitation techniques to each application domain is based on the promise to optimize validity of results by taking into consideration intrinsic characteristics of each application domain, expectations and goals among various stakeholders and diverse requirements prioritizations [5-7].

From this perspective, in this paper we reflect on our experiences in user requirements elicitation of an innovative mixed pipeline for assets creation within the video game application domain. Video game development has become over the years a collaboration activity which embraces complex communication and collaboration processes among several stakeholders who

share different roles within the video game development pipeline [8,9]. As such, it represents a unique domain in which requirements elicitation embraces challenges like: (a) creating and sharing a common understanding among technology developers and UX experts related to the underlying technology and its effect on current workflows and procedures; (b) define a custom requirements validation framework bootstrapped on this particular case study; (c) recruit representative end-users in order to perform the requirements elicitation study; (d) communicate in an efficient way the vision to end-users aiming to elicit their views, opinions, motivation, concerns and requirements related to the envisioned approach; and (e) consolidate results of the requirements elicitation with technology partners aiming to find the best trade-off solution between user wishes and feasibility.

In the aforementioned context, a three phase requirements elicitation method was adopted, as depicted in Figure 1:

- *Phase A-Preparation,* aimed to transform the vision statement of the new approach into a set of presentable usage scenarios;
- *Phase B-Requirements Elicitation*, verified the view against the perceptions and opinions of real game developers; and
- *Phase C-Analysis and Consolidation*, analyzed findings of the previous step, consolidated and prioritized results with partners and finally documented the final end-user requirements.

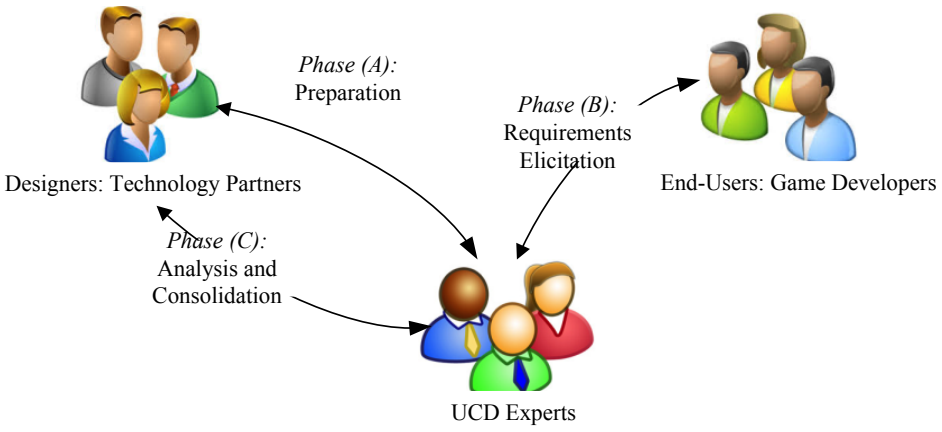


Figure 1. Overall requirements elicitation methodology.

The adopted method aimed at addressing the aforementioned challenges and reaching an optimized trade-off solution among end-users requirements and constraints related to feasibility, available time frames and resources. The rest of this paper is structured as follows. We first provide background information related to the context of the study and present the rational and

motivation behind the vision to design an innovative pipeline for assets creation in video games. Consequently we present each step of the method that was adopted with the aim to identify and validate the user requirements in the aforementioned context. Finally, we present lessons learned and summarize our findings and conclude the paper.

2 Context of the Study

According to [10], art, design and programming accounted for nearly half of the total retail cost of a next generation video game, while the remainder went to marketing, distribution and retail mark-up. In this realm, there is profound need to propose an innovative development pipeline for the creation of video game assets in order to drastically reduce both time and expenses involved in their creation, and make high quality realistic contents accessible even to small game developers [20]. From a technology point of view, this will be achieved by means of new image-and video-based technologies [11, 12] that will be developed and optimized and smoothly integrated into most dominant video game development engines (e.g. Unity3D, Unreal). From a procedural point of view, game developers will be able to capture and reconstruct real life objects by simply taking a few sequences of photos and videos to be processed by a semi-automatic software using image based rendering techniques integrated with traditional assets made of polygons and textures. As mentioned in [20]: "*It is anticipated that the new approach will significantly affect the content creation pipeline in video game development with new tools which will allow a much faster turnaround time from the idea to the prototype implementation of video games*".

The envisioned pipeline is comprised of the following major steps as depicted in Figure 2: a) the capturing step; b) the reconstruction; and c) the edit and play step. In particular, with regards to the capture step, the idea is to allow representative users to take images and videos of real life elements (e.g. buildings, houses, streets, cars, moving trees) and reconstruct them through innovative image and video based rendering algorithms and subsequently use them within a video game engine with the aim to create video games.

Figure 2. Major steps in the proposed pipeline (*source: www.cr-play.eu*).

3 Phase (A) – Preparing the user requirements study

The main objectives were: a) to perform a stakeholder analysis with the aim to understand stakeholder values, motivation and concerns in adopting the new pipeline (i.e. anticipated impact that the system would have within their current working context); b) to elaborate representative usage scenarios and to identify primary and secondary users for the tools and c) to define a tailor made requirements elicitation method based on the aforementioned tasks.

For achieving the aforementioned goals several known techniques were applied like stakeholder analysis, current workflow analysis and user scenarios. In Table 1 we summarize the scope, the method and main results of each applied technique.

Technique	Main contribution to the requirements elicitation framework
Stakeholder analysis	*Scope:* To identify stakeholder categories and understand their values, motivation and concerns in adopting the new pipeline. Finally, to define the anticipated impact of the new pipeline in their current working context and activities. *Method:* Literature review, interviews, group discussions and consolidation with consortium partners. *Results:* Stakeholders were group in two major categories, direct and indirect involved stakeholders (Appendix A). The analysis of the anticipated impact revealed several UX dimensions which were considered as important investigating within the subsequent study.

	With regards to its indirectly affected stakeholders, following dimensions were identified: a) the intention of adopting the mixed pipeline for video game development with an emphasis on validating the usefulness in current and future projects; and b) the impact on achieving the business goals and the possible effect in internal organizational or operational structures. With regards to directly affected stakeholders, following dimensions were identified: a) the perceived ease of use of the tools; b) the completeness of gathered end-user requirements with respect to the identified tools, functionalities and workflows; and c) the perceived usefulness.
Current workflow analysis for assets creation in video games	*Scope:* To identify, based on the aforementioned stakeholder analysis, primary and secondary users of the proposed pipeline. Furthermore, to specify in detail how this new approach will affect the traditional workflow of assets creation in video games. In particular, questions that were investigated were related to: a) where and when the interaction takes place; b) who is interacting with the technology; c) how the user is interacting in terms user input, what does the system do, what results does the system provide to the user; and finally d) when the user will be satisfied with the obtained results. *Method:* Literature review, interviews and consolidation with consortium partners. *Results:* The traditional workflow analysis contributed significantly in identifying the primary and secondary users of the envisioned tools (Appendix B). The technical specialists group, mainly artists and programmers, represent the primary user category which will directly interact with the interactive tools. However, there are also secondary users (the art directors, game designers, game producers, publishers and video game players) who are implicitly related to the results of the proposed pipeline. It also revealed several layers of abstractions related to the pipeline which were: system implementation layers (*system layer*), the available tools (*tool layer*) aligned to specific user category, user goals (*user layer*) and business objectives (*business layer*).

Usage Scenario Description	*Scope:* To elaborate a representative use case scenario with the aim to present it to end-users in the subsequent step. The objective of the usage scenario description was to highlight user interaction aspects of the new pipeline in terms of workflows, information architecture, presentation, functionalities and data formats. *Method:* Interviews, group discussion and consolidation with consortium partners. *Results:* The scenario description that has been elaborated consisted of several layers namely: a) the user layer; b) the tool/goal layer; and c) the activity layer (Appendix C). Based on this approach, we elaborated a representative usage scenario for each step of the pipeline (i.e. capture, reconstruct and play) along with the pre-requisites related to human and technology factors (e.g. which software or hardware would be necessary, specific human skills, knowledge or previous experience) aiming to achieve a certain goal.

Table 1. Applied techniques during phase (A).

4 Phase (B) – Requirements Elicitation Study

The objective of this phase was to capture requirements for the pipeline with the involvement of end-users, i.e., video game developers. For this purpose, we recruited (8) game developing companies, all of small medium size, from (3) different countries in the EU: Italy, Finland and Greece, and we conducted (N=17) semi-structured interviews with different stakeholders. Prior the interview the participants followed a detailed presentation related to the usage scenario that has been elaborated in the previous phase of the applied methodology.

As shown in Table 2, the participants shared different roles within the game content creation pipeline and stated that they were experienced professionals. It should be mentioned that 15 participants are usually enough to provide a solid ground for analysis [13,14], while some suggest that even smaller number is enough in case of experts [15,16].

Role	Total Number	Average Years of Professional Experience	Average Age

Producer	2	14	37
Game Designer	6	4.4	31.1
3D Artist	5	6.4	28.6
Game Programmer	4	9.7	40

Table 2. Participants' information.

Furthermore, based on findings of the previous phase, we created a tailor made requirements elicitation framework. A high level presentation of this framework is shown in Figure 3, which aligned specific dimensions with stakeholder groups.

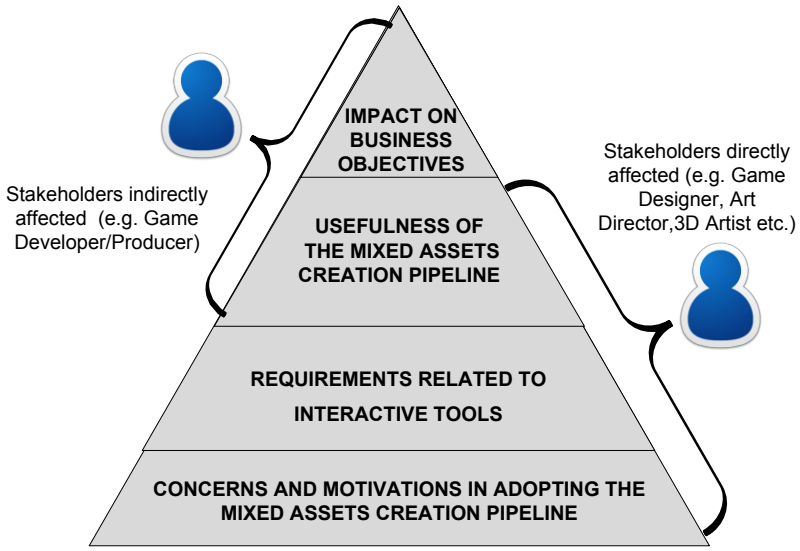


Figure 3. Requirements analysis perspectives.

The framework embraced the identified major stakeholder categories and investigated following perspectives: the adoption of such new perspectives focusing on values, motivations and intentions of end-users, the completeness of requirements with regards to the new tools focusing on scenarios as expressed in the usage scenario description, and the impact of the pipeline in business objectives. Based on the aforementioned framework, semi-structured interviews were prepared in order to determine the perceived values and difficulties on creating intentions, actual using the system and measuring the CR-PLAY impact in current processes and workflows with regards to video games content creation. In particular, the interviews were applied in order to determine the perceived values and difficulties on creating intentions, actually using the system and measuring the impact in current processes and workflows with regards to video games content creation. The interview questions were open ended aiming to allow the interviewee to express freely

her/his opinion related to several issues that were investigated during the study (e.g. *Please describe your main concerns with regards to the capturing activity as described (initializing, capturing, guidance, post processing steps, preferred output etc.)*

As shown in Figure 4, the method followed in the requirements elicitation phase was split into three steps: i) to capture the game developers' current content creation pipeline, workflows and procedures; ii) to present the proposed mixed pipeline to game developers, and iii) to analyze their views and perceptions through semi-structured interviews.

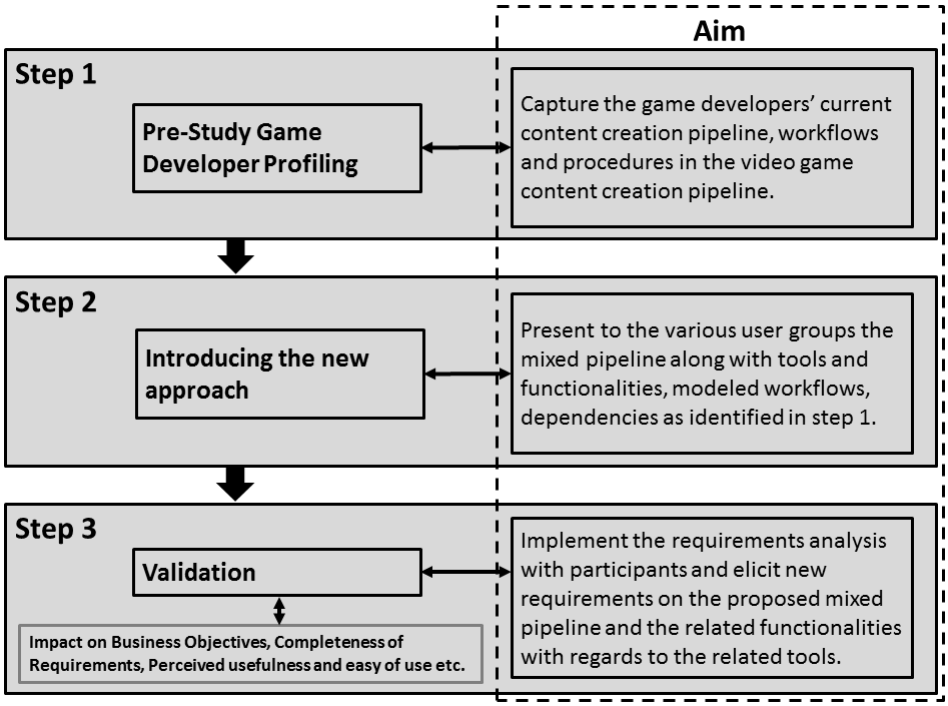


Figure 4. Requirements analysis implementation method.

5 Phase (C) – Analysis of Results

The data analysis phase of the requirement elicitation study followed sequential steps. First, the answers from all semi-structured interviews were transcribed by the UX team. For the transcription we adopted a coding schema based on the identified validation dimensions: a) impact on business objectives; b) perceived usefulness of the pipeline; c) requirements completeness focusing on functional and non-functional requirements and d) intentions, concerns and motivations in adopting the proposed pipeline.

We coded the participants' responses by reflecting the validation goals and annotated the participant responses according to the discussion theme, the

game developer company, the participants' role and the unique identifier of the interview. Following this approach we created a common template that was used by the analysts group. An example of the interviews' transcriptions is shown in Figure 4 where (A) indicates the validation goal, (B) indicates the participants' transcript label (*indicated the time of the audio recorded interview the participant expressed his opinion followed by his companies abbreviation, his role within the company and the interview unique identifier*), (C) indicates the interviewee's response and finally (D) correlates the participants response with the validation question.

A	Perceived usefulness	Motivation	C
B	[18:25] {TL_3DARTIST_ID:1}	*"Main thing is to have QUICK, PHOTOREALISTIC ENVIRONMENTS FOR OUR GAMES"*	D

Figure 5. Semi structured interviews transcript example.

The transcriptions were added on a common template and were discussed with the analysis team. Subsequently, the results were analyzed by summarizing the prevalence of categories and identifying further groupings or relationships through brainstorming sessions which were conducted by the analysis team. In addition, various content analysis techniques, such as frequencies or counts of events/mentions were performed along with narrative and correspondence analysis that aimed to create user role profiles in accordance to their responses to the semi-structured interviews.

The interviewed data were analyzed by a team that consisted of 5 UX experts. Initially, each team analyzed the data separately according to the aforementioned coding. Then, the teams arranged a round table discussion aiming to discuss findings derived from the conducted interviews and plan further activities. Next, the end-user requirements were grouped according to the requirements framework. Finally, a draft end-users requirements analysis report was consolidated with the technology partners. This task was necessary since that the requirements elicitation embraced several requirements as expressed by end-users which are from a technological point of view difficult to implement or even out of scope of the project.

6 Discussion and Lessons Learned

We summarize below our main experiences and lessons learned towards achieving the main goals of the requirements elicitation process.

Phase (A)- Preparation. Objectives: Creating and sharing a common understanding among technology developers and UX experts with regards to

the underlying technology and its effect on current workflows and procedures. Define representative usage scenarios for end-users. Define a custom requirements validation framework bootstrapped on this particular case study. Recruit representative end-users in order to perform the requirements analysis study (2 months time).

The elaboration of a representative usage scenario derived from the need to present to end-users a good real life example on how this new approach would affect their everyday activities with the utter aim to elicit requirements. We underpin that several iterations among UX experts and technology partners were required for achieving this task. This iterative cycle was primarily necessary in order to create and share a common understanding among different technology partners and UX experts within the consortium related to user interaction perspectives. We also stress, that the active and collaborative involvement of technology partners in various techniques during phase (A) facilitated reaching a common ground and mutual understanding related to the involved steps of the pipeline aligned with stakeholder categories, user roles and activities.

In particular, the stakeholder analysis contributed in the identification of the effects of the envisioned approach to each stakeholder category and separated the impact of the envisioned technology to multiple levels of abstraction per stakeholder group (e.g. business, operational, organizational and procedural). Subsequently, the traditional workflow analysis related to video games assets development contributed significantly in further identifying the primary and secondary users of the envisioned tools. It also, revealed the main procedural impact of the new pipeline and analyzed interactions in several levels like system procedures (system layer), functionalities as exposed by the tools of the new pipeline (tool layer), user goals (user layer) and high level objectives (business, organization, operational layer).

Finally, the usage scenario elaboration schema (Appendix C) was perceived as very helpful by the technology partners in order to formalize the user interaction activities. The schema inquired specific information related to the pre-requisites of representative related to human and technology factors and details for the anticipated human computer interaction in each step of the pipeline like where and when the interaction takes place, who is interacting with the technology, how the user is interacting in terms user input, what does the system do, what results does the system provide to the user, and finally when the user will be satisfied with the obtained results.

It is also important to mention that the aforementioned approach revealed the need to identify a specific requirements elicitation framework with the aim to focus not only in user interaction tasks but also to take into considerations

other aspects which also affects stakeholder groups. Accordingly, specific instruments were developed (questionnaires, semi-structured interviews and focus group studies discussion topics) aiming to address specific functional, non-functional requirements, intentions of use and end-users motivation and concerns.

One more important aspect was related to recruiting representative end-users. Given that the involvement of real end-users is a very critical aspect with respect of the internal validity of a study (since it determines in a high percentage the accuracy of delivered results) it is important to consider the user recruitment, even at a project proposal stage, as one of the most critical tasks and contain contingency plans aiming to resist in last minute cancelations of end-users.

Phase (B) – Implementation. Objectives: Communicating in an efficient way the vision to real end-users the new approach aiming to understand their views, opinions, motivation, concerns and requirements related to the envisioned approach. (1 month time)

The implementation of the study was performed smoothly, despite that the requirements elicitation needed to be implemented within a multinational context and that innovative aspects needed to be presented in a clear and understandable way to end-users. From a procedural point of view we decided to translate the initial presentation for end-users and semi-structured interviews in the local language of each game developer aiming to present the pipeline in a more efficient way as well to non-proficient English speakers. Based on the same rational, the interviews were performed by consortium partners in their local language and were afterwards translated to English. The participants followed an initial presentation prior the conducted interviews. Furthermore, the interview structures with the representative scenario were sent to the interviewees before the interview appointments. In each requirement validation session conducted, audio was recorded (with the consent of the participants) and took on average 1 to 1 and 1/2 hours.

The new technology along with its vision was presented in different levels of abstraction based on the adopted framework (e.g. technical, procedural, conceptual and business level) to seventeen (17) persons who had different roles in the video game production pipeline and were end-users. Participants were asked to express their thoughts and were encouraged to be as precise as possible and were asked not to hesitate to provide positive or negative feedback on the themes of discussion. Focus group studies were also organized, after the semi-interviews, in order to triangulate results and

facilitate the exchange of ideas and enhance discussions among participants who shared different role within the video game development pipeline.

Phase (C) – Analysis and Consolidation. Objectives: Consolidate results of the requirements analysis aiming to find the best trade-off solution between user wishes and feasibility. (1 month time)

The participants of the requirements validation study share a positive attitude towards adopting the proposed mixed development pipeline of contents creation. They perceived the approach as very useful and innovative with the potential of solving an outstanding issue. However, besides the positive predisposition expressed by the game developer companies in adopting the proposed approach, the requirements validation study revealed concerns. The critical concerns of the participants are related to the quality of the representations, the degree in which the representations can be modified, the capability of controlling the quality of the assets as well as support for team member communication during the capture phase, the guidance during the capture that must be related with the desired quality of an asset and the overall management of the captured and reconstructed assets within a certain game design. Several functional and non-functional requirements were identified which are further presented in [19].

One important aspect from the analysis phase is that some of the end-users resisted to change and maintained a bias towards their current way of thinking and acting, in terms of methods, tools used. So in case of disruptive new technologies, like the proposed one, they may fail to envisage the new potential and capabilities offered, that may change dramatically the way they conduct their work. However, the same attitude was recognized, in some cases, by technology partners since they resisted in adopting end-users views, opinions and expressed requirements. In addition, some end-users often felt threatened by the new approach, as their skills (e.g. in the case of 3D artists and modelers who will be affected by this new approach in several ways) may become useless and obsolete. In particular, some participants noticed: "*Adopting the proposed approach means that we will not do any modeling or we will model the basic structure. Seems that we are not needed. [3D artist]*".

7 Conclusions

The work presented in this paper describes a case study for requirements elicitation. In particular, the aim was to identify and prioritize requirements for an innovative assets creation pipeline for video games development. As such, the main contribution of this paper is to present a body of knowledge, based

on our experience, which may be of interest by UX researchers and practitioners who are facing similar challenges. From this perspective, the paper presents a case study during which stakeholders perceptions, goals and needs have been captured with regards to an envisioned new approach for assets creation during video games development. These requirements are an essential first step of the innovative tools development cycle and were gathered by adopting an iterative process. In the context of the reported study, this process involved the active participation of the technology providers, who developed the capturing tools and the image - video based rendering technology. The tensions between the technology affordances and the users' requirements were identified and discussed in successive stakeholders' analysis meetings, an important part of the requirements analysis framework.

This research could be also of importance for the video game industry and game players' experience since it discusses the introduction of advanced technologies in the form of tools to be used by the game designers and developers themselves. Study of the process of designing these tools, and the way these are introduced affecting the way new generations of games are designed in the future, is also important for the game industry.

Acknowledgements

This work is supported by the European Seventh Framework Program under grant ICT-611089-CR-PLAY.

Appendix A. Stakeholder Analysis

Stakeholder Category	**Role in the video game creation pipeline**	**Anticipated Impact in current activities**
Game Producer / Game Developer	Handle all development costs set and maintain deadlines and the overall budget. Overview the progress of the video game development tasks to assure quality and timely delivered outcomes.	A **significant competitive edge** to the European SME's in the game sector.
Game Publisher	Handles all publishing costs (materials production, advertising, etc.) Coordinates with press, plans events at	An important **positive impact** on time-to-market.

	gaming conventions (etc.).	
Game Designer	Game designers originate ideas for new games, determining the game mechanics, the users motivation etc.. They create a detailed design document to guide the various art and technical teams executing the game. Usually a design document specifies story, game play, settings, characters, environments etc.	More creative games can be produced within shorter time. An important **positive impact** on **development costs** and **time-to-market**.
Art Director, Concept Artists, 3D Animators, Engineers	Run the entire creative process, providing direction and feedback to the various teams of artists in the content creation pipeline.	A **major shift** in the content creation pipeline in terms of **simplicity, speed** and **quality**.
Capture Expert	Capture real life assets according to a certain game design and review the quality of the proposed result according to a specific game design	*Not applicable since this is a currently non-existing stakeholder category*
Video Game Player	This category represents the final users of the video game development pipeline. The players are the final "customers" the video game industry is targeting to.	Through the proposed approach more **creative** video games will be available for game players faster and cheaper.

Appendix B. Current Workflow Analysis

The figure below depicts the main human roles and summarizes associates workflows in the video game asset creation pipeline [17, 18]:

(a) The *game designer* in collaboration with others (*i.e. the game producer, the art director, etc.*) envisions the entire game and creates a detailed design document.

(b) The *concept artist* is usually prototyping according to the game design document various characters, environments and objects.

(c) The *modeler* responsibility is to take a two dimensional piece of concept art and translate it into a 3D model that can be given to animators. There are a variety of 3D modelers depending on the specific role: environmental, character or objects modelers.
(d) The *technical engineers* ensure that the game assets are easily integrated into the game engine and are behaving as stated in the game design document. They are also responsible for balancing the quality according to technical limits of the chosen platform (*e.g. depending on the platform, mobile or web, the technical engineers decide about high or low rendering etc.*). The technical team tasks include among others to implement the game mechanics and to proceed with texturing and lighting tasks etc.
(e) – (f) Several iterations among the art director and aforementioned team members are important in order to assure high quality results.

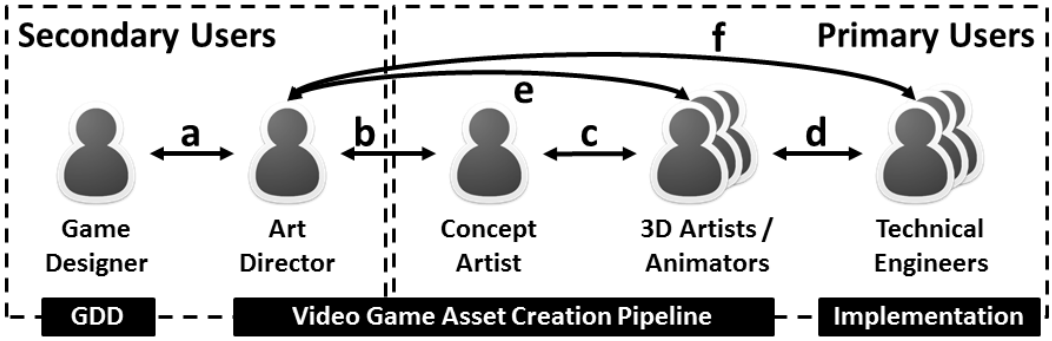


Figure 6. Current workflow analysis.

Appendix C. Detailed Usage Scenario Description

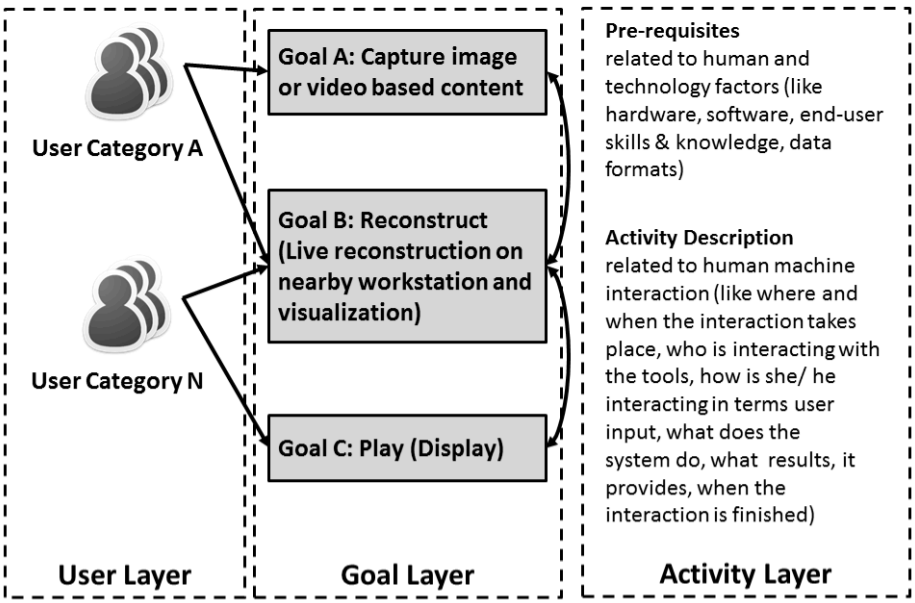


Figure 7. Usage scenario elaboration schema.

Usage Scenario:
Paolo wants to make use of a mixed pipeline for content creation to create a prototype of the game set in his hometown. He urges the team's programmers and game designers to familiarize themselves with the tools and technologies developed. The production of the game prototype is organized so as to re-use animations, tricks and features of the game together with photorealistic environments and objects captured and displayed within the mixed pipeline. In particular there is great interest in having ***high quality environments and animated elements (fountains, moving trees, flags, moving cars etc.)****, captured from the real world and mixed with characters already created with a traditional approach. Another important feature for the game is the possibility to* ***represent different moments of the day*** *thanks to the* ***dynamic lighting model*** *supported by system*."

Paolo has several goals when using the pipeline:
User Activity (a): Capture image or video based content

Pre-requisites: Paolo is using the capturing tool on his new Smartphone device (iPhone 5). He downloaded the app from the apple store. The mobile device has a high resolution camera and contains gyroscopes and inertial sensors which help estimate small camera-phone pose changes. To capturing details of surface normals especially for single objects (which is a very demanding task) he knows that a specialized setup is required (e.g. the mobile camera would sit in a lightweight cradle and it would capture image sequences while triggering LED lights, attached to the four sides of the cradle, i.e. above, below, left, and right of the camera).

Goal - Initializing: Paolo activates the capture tool. The tool provides Paolo with several functionalities. Paolo gives some basic information about the scene structure (e.g. indoor / outdoor, single object / complete street) the system will decide how many images are needed, and it will provide guidance on the positions they have to be taken from. Furthermore, the capture tool allows user-based annotations and extends them to the complete scene for a better segmentation/separation of objects.
Goal - Capture: The system will decide automatically if enough data has been captured or will guide the user towards new camera positions that support a more complete reconstruction. The user is then guided towards the new view, and the image is captured automatically at the right position. In order to actually capture a computed view, the user needs to move the device into the correct position. The feature matching and pose estimation of the next best view (NBV) is done constantly in about 1 second but the global optimization gets more expensive when the number of images and 3D points increases.
Goal - Guidance: The capture tool uses the phone display to provide guiding to Paolo in two ways: First, it presents an estimate of how the new image has to look like, in order to give a global impression of the camera position and viewing angle. Second, it shows an approximation of the direction towards the exact 3D position. Visual aids will be provided, such as arrows guiding the user to the NBV.
Goal - Iterations: The next iteration starts after the internal reconstruction has been updated. The capture application registers all captured images, creates a sparse reconstruction, and uses this data to estimate new views and to guide the capture process.
Goal - Post processing: Paolo has finished capture the desired scene. For the quality of a final reconstruction it can also be important to exclude captured data if it introduces too much uncertainty. As a consequence Paolo is using on a nearby workstation a system that post-processes data and removes unnecessary or outlier data before a complex reconstruction. It can also be important to separate a dataset into several smaller parts. In fact this activity develops confidence measures as a post processing step, and improved reconstruction, which, in contrast to previous techniques will concentrate on providing the 3D information useful for IBR and VBR, rather than an accurate-as-possible geometric model.
Results: Paolo has finished the post-processing step. He can now proceed with the following actions aiming to manage captured indoor or outdoor scenes. The available options are: a) Review which provides him a real time feedback about the captured scenes, he actually can move to different viewpoints of the captured scene, b) Delete one or all images/videos takes in this activity, c) Save the capture activity, d) Share/Send to other members in the content creation pipeline, e) Edit and finally f) export as external files. The export functionality allows Paolo to move the captured content to the reconstruction stage.

User Activity (b): Reconstruct (Live reconstruction on nearby workstation and visualization of the resolution)

Pre-requisites: Paolo has finished the capture activity of a scene and he has exported in a pre-defined IBR/VBR format the following data: a) multiple photos of a scene, b) calibrated cameras for each photo (i.e. the position and orientation of each photo), c) 3D point cloud with normals (*usually sparse and often inaccurate*), d) a mesh, usually approximate, can contain holes and inaccuracies. Other data may also be included. All this information is saved in an IBR/VBR format aiming to be integrated to the reconstruction tool.

Goal - Import files: Paolo uses the reconstruction tool to import the captured data. It is a stand alone application that is used for reconstructing the captured scene. As seen in the figure below tools provide a coarse (sometimes inaccurate) 3D version "point cloud" or mesh.

Goal - Specifying preferred resolution: Through the reconstruction tool offered functionalities the user can choose among high or low resolution representation of the captured scene, Depending on the produced quality the end-user can decide whether the capture step has been accomplished successfully or whether he needs to re-engage in the capture activity.

Results: The output of the reconstruction tool is 3D information (point cloud, normals, mesh), which can be incomplete and inaccurate as they are comprised from a collection of images/videos taken from known camera viewpoints. Paolo uses his laptop aiming to review – evaluate the produced outcome and eventually start over with the capture process or re-capture certain view positions. This will be done by viewing the result with the renderer / play stage. The additional cost to see the result with IBR is minimal. The functionalities provided by the PLAY tool allow Paolo to navigate through the produced model aiming to examine continuity and diverse points of view with regards to the produced 3D model of the captured backdrop. Given that Paolo has the game design in mind he can evaluate whether the produced results are appropriate

User Activity (c): Play (Display)

Pre-requisites: The reconstruction step has finished and now it is the time that the resulting IBR/VBR representations are displayed in a game engine environment (e.g. Unity 3D).

Goal - Display the 3D model: The captured and reconstructed environment is displayed through an IBR/VBR format for display within the game engine.

Goal - Play: The usability dimension in this approach relies on the fact that Paolo can use, as he did, all the available functionalities of his favorite game engine in a unified manner with regards the virtual and real world captured objects. Among others the editing tools which will be embedded within a standard game engine will support dynamic lighting / delighting

modeling, will allow high quality rendering of the imported assets, high-quality real-time in-painting (*allowing users to move around more freely in captured environments, even when part of the output image includes regions not seen in the input photographs*) and will allow class-specific surface synthesis. With regards to imported 3D video based data, artists will be able to encode, as a view-dependent Video Sprite, effects such as sand or snow being kicked up with foot-falls, or dirt being sprayed by spinning wheels. During play, the data will be rendered to match the environment and new viewing angle, and when the action calls for it. Here, the game designers will be able to specify which of the recorded dust-clouds is triggered by a big vs. a little impact.

Goal - Rendering: Through the IBR / VBR plug-ins the rendering contains many photo-realistic features such as shading, texture, reflection, shadowing, motion blur, transparency and depth of field -- depending on the type of capture – creating a lifelike perception.

Results: Paolo can use the photorealistic environments and objects captured together with characters and assets already created with a traditional approach. The game engine handles both types of content in a unified manner therefore Paolo can continue to develop the game mechanics of the game etc.

References

[1] Silva da Silva, T., Martin, A., Maurer, F., Silveira, M. (2011). User-Centered Design and Agile Methods: A Systematic Review. In: Proc. of the Agile Conference (AGILE 2011), pp. 77-86

[2] Rebaque-Rivas, P., Gil-Rodríguez, E., Manresa-Mallol, I. (2010). Mobile learning scenarios from a UCD perspective. In Proc. of the Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '10), pp. 389-390.

[3] Asuncion, H. Socha, D., Sung, K., Berfield, S., Gregory, W. (2011). Serious game development as an iterative user-centered agile software project. In: Proc. of the International Workshop on Games and Software Engineering (GAS '11). ACM, pp. 44-47.

[4] Cremers, A., Jong, J., Balken, J. (2008). User-Centered Design with Illiterate Persons: The Case of the ATM User Interface. In: Proc. of the International conference on Computers Helping People with Special Needs (ICCHP '08), Springer-Verlag, pp. 713-720.

[5] Pardo, S., Vetere, F., Howard, S. (2005). Broadening stakeholder involvement in UCD: designers' perspectives on child-centred design. In: Proc. of the Australia conference on Computer-Human Interaction: Citizens Online: Considerations for Today and the Future (OZCHI '05). Computer-Human Interaction Special Interest Group (CHISIG) of Australia, pp. 1-9.

[6] Popa, D. (2013). Industry Design Case: Introducing Gamification Persona Tool. In: Proc. of the Conference on Human Factors in Computing Systems (CHI 2013).
[7] Yiing Y'ng, N., Chee Weng, K. (2014). A review of affective user-centered design for video games. In Proc. of the International Conference on User Science and Engineering (i-USEr), pp.79-84
[8] Pagulayan, R. J., Keeker, K., Wixon, D., Romero, R. L., & Fuller, T. (2002). User-centered design in games. The HCI handbook: Fundamentals, evolving technologies and emerging applications (pp. 883-906). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
[9] Asuncion, H. Socha, D., Sung, K., Berfield, S., Gregory, W. (2011). Serious game development as an iterative user-centered agile software project. In: Proc. of the International Workshop on Games and Software Engineering (GAS '11). ACM, 44-47.
[10] Wesley D., G. Barczak 2010. Innovation and Marketing in the Video Game Industry: Avoiding the Performance. Chap. 10. Surrey, UK: Gower, 2010.
[11] Laffont, P.-Y., Bousseau, A., Paris, S., Durand, F., and Drettakis, G. 2012. Coherent intrinsic images from photo collections. ACM Trans. Graph. 31, 6 (Nov.), 202:1–202:11.
[12] Ballan, L., Brostow, G., Puwein, J., and Pollefeys, M. 2010. Unstructured video-based rendering: Interactive exploration of casually captured videos. ACM Trans. Graph. (Proc. SIGGRAPH) 29, 3, 87:1–87:11.
[13] Miles, M. B., Huberman, A. M., & Saldan_a, J. (2014). Qualitative data analysis: a methods sourcebook. Thousand Oaks, Califorinia: SAGE Publications, Inc.
[14] Guest, G.; Bunce, A. & Johnson, L. (2006). "How many interviews are enough? An experiment with data saturation and variability". Field Methods, 18(1), 59-82.
[15] Morse, J. M. (1995). The significance of saturation. Qualitative Health Research, 5(3), 147-149.
[16] Morse, J. M. (2000). Determining sample size. Qualitative Health Research, 10(1), 3-5.
[17] Bates, Bob (2004). Game Design (2nd ed.). Thomson Course Technology. ISBN 1-59200-493-8.
[18] Bethke, Erik (2003). Game development and production. Texas: Wordware Publishing, Inc. ISBN 1-55622-951-8.
[19] Christos Fidas, Jan Storgårds, Melpomeni Masoura, Nikolaos Avouris (2014). Deliverable 4.1 End-user Requirements Analysis. Public available: http://www.cr-play.eu/deliverables/
[20] Online resource: http://www.cr-play.eu

Solution-based Requirements Capture with PDot in an E-Learning Context

Matthias Heintz, Effie Law
University of Leicester,
University Road, Leicester, LE1 7RH, United Kingdom
mmh21@leicester.ac.uk, lcl9@le.ac.uk

Abstract. This paper presents the requirements engineering work conducted in the Go-Lab project, which aims to improve science education by enabling teachers to deploy online labs and scaffolding apps to enrich the traditional classroom-based teaching and learning. We tackled the challenge of developing a Participatory Design approach to address the specific arrangement of the project where the initial user-based requirements engineering process and the technical implementation process were launched at about the same time instead of running the two processes one after the other. To support the PD activities, PDot, an online tool to gather user requirements and feedback, has been developed. In this paper, we present the design of PDot together with preliminary usability and User Experience evaluation results and future development plans. By presenting our challenges and approaches we aim to stimulate discussions about our proposed solution and possible improvements and to inspire other research projects with similar setups.

1 Introduction

The Go-Lab (Global Online Science Labs Inquiry Learning at School) project[1] is a European initiative to support science teaching by enabling teachers to incorporate online resources in science lessons [2], especially online labs and scaffolding apps. The requirement engineering (RE) work described in this paper has been conducted in the context of Go-Lab with the aim of facilitating its software design and development process. Go-Lab supports three major types of online labs, namely, remote labs, virtual labs and datasets. Remote labs are real laboratories that are controlled through a website. An example is WebLab-DEUSTO Aquarium[2], where real balls with different density can be thrown into tubes filled with different liquids using a web interface, and results can be observed through webcams to learn about Archimedes' principle. Virtual labs are simulations that emulate real labs. An example is

1 http://www.go-lab-project.eu/ [last accessed: 22/06/2015]

2 http://www.golabz.eu/lab/weblab-deusto-aquarium-0 [last accessed: 22/06/2015]

Splash: Virtual Buoyancy Laboratory[3], where virtual balls with different densities can be created and then thrown into virtual tubes filled with different liquids. Datasets represent the results of experiments already conducted, for example, with the Large Hadron Collider at CERN. An analysis tool is then used as online lab to visualize and work with those results. An example for this approach is HYPATIA[4]. Although in this case the real lab is not controlled through a web interface that allows users to conduct the experiment at the spot, the results can be used as if they would have been obtained that way.

By scaffolding apps, we refer to applications that facilitate learning such as the Hypothesis Scratchpad[5], which enables learners to create hypotheses by drag&drop, using predefined building blocks with words and expressions given by their teacher and with their own terms.

The Go-Lab project develops new and innovative online labs and scaffolding apps as well as collecting existing ones in a repository website [5]. To ensure compliance of the developed digital artefacts with the requirements and visions of the prospective users, a Participatory Design (PD) approach [11] has been applied in the project.

2 Motivation

In the Go-Lab project, the prototype development phase was kicked off more or less at the same time as the requirements engineering (RE) process involving end-users. This setup was somewhat unusual, because typically a user-based RE process takes place, at least in the initial project phase, prior to implementation. However, in the context of Go-Lab, some mock-ups of the main features were created based on the legacy of the related projects. To some extent this could be considered as advantageous, given that bootstrapping a new system can be very challenging and time-consuming, and it is always advisable to avoid reinventing wheels. Adapting existing designs with the ultimate goal of enhancing them can be more cost-effective than starting the design from scratch, especially when potential end-users may have limited or even no experience with online labs. Nevertheless, ongoing

---

3 http://www.golabz.eu/lab/splash-virtual-buoyancy-laboratory [last accessed: 22/06/2015]

4 Hybrid Pupils' Analysis Tool for Interactions in ATLAS (http://www.golabz.eu/lab/hypatia-hybrid-pupils-analysis-tool-interactions-atlas [last accessed: 22/06/2015])

5 http://www.golabz.eu/app/hypothesis-tool [last accessed: 22/06/2015]

involvement of end-users to improve the proposed design is indispensable. This complies with the principle of Participatory Design (PD), whereby end-users can exert direct influence on the proposed design by voicing out their needs or preferences and by articulating their ideas that will seriously be considered and addressed during the development process.

Feedback can be collected in the form of solutions proposed directly by users to address undesirable, unusable or missing features that they experience or discover when interacting with the prototype. This is alternative to the conventional waterfall-like approach: first formally specifying requirements based on users' input and then asking experts to evaluate the implemented solution against the specification. We argue that this **solution-based** requirements specification can generally be more actionable and faster to implement than the traditional way of translating users' needs/requirements into design solutions, whereby mitigating the issue of being lost in translation [3] and of the filtering effect [10].

Requirements engineering and prototype development are normally conducted in tandem as well as in iteration. This entails user-based evaluations of increasingly interactive prototypes, ranging from non-interactive mock-ups to executable applications with most of the functionalities required. To accommodate this need, traditional Participatory Design methods, which are mostly paper-based, are deemed inadequate. For instance, the interactivity of an executable prototype cannot be evaluated with static paper-based screenshots.

As compared with the paper-based approach, an online tool can effectively support the evaluation of interactive prototypes. Putting in the context of Go-Lab, existing applications have been evaluated in terms of their compatibility with the goals and objectives of Go-Lab. In particular, with partners of a European project geographically distributed, it is necessary as well as desirable to have a tool that can address this distance barrier, considering potential advantages in face-to-face as well as distributed situations [8]. Nonetheless, as no appropriate technical solution is yet available, we have been motivated to design and implement our own, which is called **Participatory Design online tool**, or in its short form, PDot [7].

Requirements have therefore been collected on two levels. On the basic level, requirements for digital artefacts developed in Go-Lab are captured, using the software-supported PD approach supported by PDot. On the meta-level, requirements for PDot itself are derived from observing users when interacting with it to provide feedback on the Go-Lab artefacts and from analysing users' responses to specific questionnaires. The focus of this paper

is the development and evaluation of PDot and proposal of the new variant of the PD method to gather solution-based UX requirements it supports.

3 Design of PDot

PDot consists of three components, as shown in Figure 1. The first component is the PDot box, which is displayed as the left hand column and is used for capturing users' feedback. The second component is the PDot Instruction Area, which is located at the upper part of the screen on the right hand side. This is used for giving guidance to the participants throughout the Participatory Design activity. The third component is the software to be evaluated, which is displayed in the lower part on the right hand side (e.g., the screenshot shown in Figure 1 shows the Concept Mapper[6] scaffolding app).

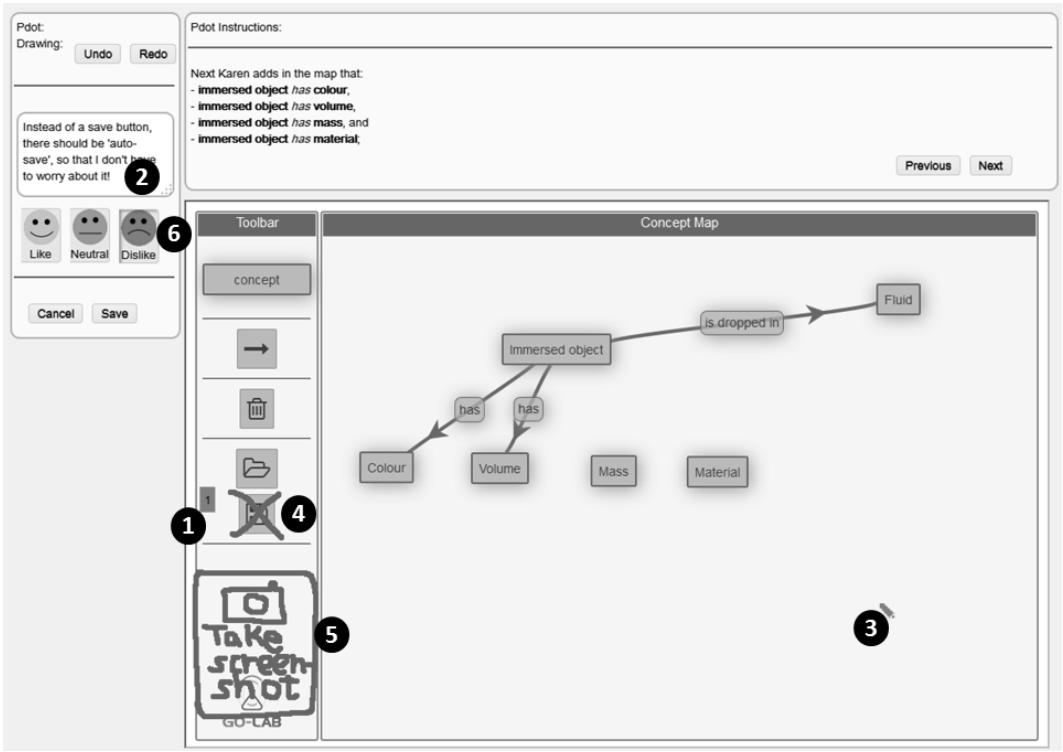


Figure 1: Screenshot of PDot (for referencing parts or elements of the picture in the text, numbers on black dots have been added)

To give feedback the participant first clicks on the position of the screen to comment on. This creates a yellow sticky note in this place (❶ in Figure 1). The feedback can then be specified textually to give an explanation and detailed description of the problem and suggested solution, using the text input

[6] http://www.golabz.eu/content/go-lab-concept-mapper [last accessed: 22/06/2015]

box in PDot (❷in Figure). At the same time the cursor changes to a blue pen (❸ in Figure), allowing the participant to additionally provide visual feedback by creating a freehand drawing using the mouse. Examples for such creative feedback are crossing out an element to suggest its removal (❹ in Figure) or drawing something, like a "Take screenshot button" (❺ in Figure) to suggest adding an element to the software. Three buttons in PDot (❻ in Figure) enable the participants to specify their mood, which can be helpful for researchers and developers later on in analysing the results.

4 Preliminary Evaluation of PDot

4.1 Procedure and Participants

PDot has been used and evaluated in several Participatory Design workshops. As the related data analysis is still ongoing, only preliminary results are reported here. To give examples from the two main user groups, teachers and students from upper primary school up to universities, the results regarding usefulness, usability and user experience from one student- and one teacher-based event are reported.

The student-based event was conducted in March 2015 with 32 first-year university students in computing who used PDot to give feedback on a complete online lesson on Electricity - An Alternative approach of Ohm's Law[7], including learning scaffolding apps and an online lab. The evaluation took place in a partly distributed setting, as the students were briefed in class but performed the actual evaluation in their own time, either at home or during another class. 28 (24 male and 4 female; all university students currently studying in the United Kingdom) of the participants filled in the questionnaire about PDot completely, 4 only partially (not answering the questions reported in this paper).

The teacher-based event was conducted in April 2015 as part of a teacher conference with 20 teachers who used PDot to give feedback on an online chemistry laboratory called BOND lab[8] on precipitations. As the tool was used as part of the workshop on site, the setting was non-distributed in this case. 19 (2 male and 17 female; 8 Primary School, 9 Secondary School, 1 Further Education college, 1 university; all from the United Kingdom) of the

---

7 http://www.golabz.eu/spaces/electricity-alternative-approach-ohms-law [last accessed: 22/06/2015]

8 http://www.golabz.eu/lab/bond [last accessed: 22/06/2015]

participants filled in the questionnaire about PDot completely, one only partially (not answering the questions reported in this paper).

4.2 Instruments

The evaluation approach we adopted was primarily subjective self-reporting. After using PDot, the participants were asked to complete a questionnaire on usability and user experience (see Appendix A). Specifically, Section 2 is based on two standardized questionnaires, namely AttrakDiff [6] and Usability Metric for User Experience (UMUX) [4], whereas Section 3 and 4 have been developed by the first author to get first impressions on the usability and feature set of PDot. Consequently the questions are based on three usability attributes: Learnability (easy to learn), Efficiency of Use (ease of use), and Subjective Satisfaction (useful) [9] enhanced with questions regarding the usefulness of different PDot features and free form fields to explain in more detail. Put concisely, AttrakDiff is grounded in the theoretical assumption that the hedonic and pragmatic quality of an interactive product contributes to its attractiveness, which in turn leads to positive user experience and intention to use. UMUX is built upon the traditional notion of usability (ISO 9241-11:1998[9]) and System Usability Scale (SUS) [1] with the aim of producing an even more parsimonious scale for industrial use.

The list below summarizes the measures taken (see Appendix A for details):

- Background
  Some demographic data have been collected from the participants, including gender, age, school type, and country. These can be co-variants for the evaluation results of PDot, but are not dealt with accordingly in this study, given the small sample size.
- User Experience (AttrakDiff and UMUX)
  These two standardized questionnaires capture quantitative data to be analysed with appropriate statistical methods.
- Usability
  A set of 11 usability statements collects quantitative data to be analysed with appropriate statistical methods.
- Features
  The open-ended questions enable participants to give qualitative comments on existing features of PDot and to propose new ones.

[9] ISO 9241-11:1998 Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability

However, as the response rate to this question is rather low, the limited findings are not reported in this paper.

5 Results and Discussions

To compare the responses of students and teachers regarding the "Usability of PDot" statements in the questionnaire a non-parametric Mann-Whitney-U test has been performed, given that the results of a Shapiro-Wilk test of normality indicates a non-normal distribution of the data. Only two question showed a significant statistical difference between teachers and students: *"Sticky notes marking the spot of my feedback were useful."* with $U=177.50$, $p<0.05$, student mean rank=20.84, teacher mean rank=28.66 and *"It was easy to switch between different modes (instructions, interact with app, give feedback)."* with $U=158.50$, $p<0.05$, student mean rank=20.16, teacher mean rank=29.66. A possible explanation why teachers (mean=3.53) found it easier, as compared with their student (mean=2.79) counterparts, to switch between PDot modes might lie in the different software artefacts evaluated by the students and teachers. The teachers were evaluating an online lab, which requires less navigation and fewer switching between interacting with the lab and giving feedback using PDot as compared to the students' evaluation of a whole online lesson, which not only includes an online lab but also several scaffolding apps and other online resources (e.g., Electrical circuit lab[10], Conclusion Tool[11] and YouTube videos). While on average both user groups found the virtual sticky notes useful, we are not sure why teachers (mean=4.32) did this significantly more than students (mean=3.71).

The other questions did not show any significant difference in the ratings between the students and teachers, suggesting that their perceptions of the functionalities of PDot are similar. For the teachers' responses, the mean values were all higher than 3.0, indicating their overall neutral attitude towards the statements or their tendency to agree with the statements. For the students' responses, the mean values were less than 3.0 for two of the statements, suggesting that they tended to disagree on them while being neutral or inclined to agree on the other statements. The first statement on which the students disagreed is *"Design of PDot was visually appealing."* (mean = 2.79, SD = 1.1) and the second one is *"It was easy to switch between different modes (instructions, interact with app, give feedback)."* (already discussed above).

---

[10] http://www.golabz.eu/lab/electrical-circuit-lab [last accessed: 22/06/2015]

[11] http://www.golabz.eu/apps/conclusion-tool [last accessed: 22/06/2015]

The rating of students for the former statement, albeit not significantly different from the teachers (mean = 3.32, SD = 1.06), indicates that the design seems sufficient for the teachers whereas the students seem to see the need for improvement. In our future work, data analysis with a bigger dataset will be performed to see if these first impressions can be confirmed.

To compare the responses of students and teachers regarding the "User Experience of PDot" statements (which were rated by 26 students and 19 teachers) in the questionnaire, an independent samples t-test has been performed for pragmatic and hedonic quality measured through AttrakDiff, given that the result of a Shapiro-Wilk normality test indicates the normal distribution of the data (significance levels greater than 0.05, see Table 1).

Table 1: Normality test results for pragmatic and hedonic quality (from AttrakDiff word pairs)

		Shapiro-Wilk		
		Statistic	**df**	***p* value**
Pragmatic Quality	**students**	.975	26	.748
	teachers	.938	19	.245
Hedonic Quality	**students**	.975	26	.755
	teachers	.907	19	.064

For the User Experience question based on AttrakDiff, no significant differences in the perception of the pragmatic quality between students and teachers have been found. But teachers rated the hedonic quality significantly higher than students did with $t[43] = -2.18$, $p<0.05$ (student: mean = -0.27, SD = 0.78 and teacher: mean = 0.37, SD = 1.19). Both mean ratings are in the average region, thus there is still room for improvement, but the teachers already perceived a higher potential for getting pleasure from the tool usage. When analysing the statements individually using a non-parametric Mann-Whitney-U test, as the results of a Shapiro-Wilk test of normality indicates a non-normal distribution of the data, two of them show significant differences. One is *"dull - captivating"* with $U=142.00$, $p<0.05$, student mean rank=18.96, teacher mean rank=28.53 and the other one is *"cheap - premium"* with $U=127.00$, $p<0.01$, student mean rank=18.38, teacher mean rank=29.32. The mean student rating of these statements goes towards the negative and the mean teacher rating towards the positive side of the scale (calculated from the 7 point scale by using values from -3 for the word on the left to +3 for the word on the right, see Table 2 for results).

Table 5: Ratings of word pairs from AttrakDiff with significant differences between students and teachers

	Students		Teachers	
	Mean	**SD**	**Mean**	**SD**
Dull - Captivating	-0.77	1.47	0.26	1.45
Cheap - Premium	-0.65	1.06	0.37	1.50

These findings correspond to the results of the analysis of the Usability questions, where the students stated that they did not find PDot visually appealing where the teachers did. For the 4 statements of the UMUX questionnaire, no significant differences in rating between the students and teachers have been found. In our future work, data analysis with a bigger dataset will be performed to see if the same trends can be found.

Analysing a bigger dataset, including teachers and students evaluating different or the same applications using PDot will also help to overcome one of the limitations of the current data analysis. As the samples reported on in this paper not only differed regarding their characteristics (teachers or students), but also regarding the artefact evaluated using PDot, the latter might also influence the results, thus limiting the certainty of the reasons for the differences. Future analysis and results will show if the responses of each target group are consistent, when using PDot to give feedback about the same and different digital artefacts, overcoming this limitation of the preliminary data analysis.

6 Future work

The preliminary results of the usability and User Experience evaluation show that usability and usefulness of PDot are perceived as sufficient, but there is still room for improvement, at least from the students' perspective especially regarding the visual representation of the tool. As PDot is currently still under development, these results are going to be taken into consideration for its future re-design and development. Besides asking usability-related questions, the questionnaire also gathered functionality-related information, which will be incorporated in the new version of PDot currently developed. For example it has been reported that it can be difficult to identify in which mode PDot currently is. This will be addressed by visually highlighting the feedback mode through a coloured frame around the prototype. An alternative idea is to "grey out" the prototype while it is not interactive, as done in other applications (e.g., the Snipping Tool in Microsoft Windows), but it has to be checked first,

if this alteration of the prototype might have an influence on the quality of feedback.

As the dataset reported on in this paper is rather small and thus only allows first impressions about the usability and User Experience rating of PDot, more data will be collected and analysed to check if these preliminary findings will be further supported.

Currently PDot allows participants to freely express their feedback and input (through freehand drawings and in a single textbox). Accordingly, the results gathered are a mixture of usability issues, feature requests (requirements), and other matters (e.g. feedback regarding the learning content). As the Go-Lab project at this stage is interested in all kinds of feedback and not only requirements, PDot provides a sufficient way to gather it for the researchers working on improving the usability and user experience of the project's results. To support researchers in the task of classifying and analysing the results gathered, (partially) automated analysis, for example, through Natural Language Processing, might be added in upcoming versions of PDot.

Acknowledgements

This work was partially funded by the European Union in the context of the Go-Lab project (Grant Agreement no. 317601) under the Information and Communication Technologies (ICT) theme of the 7th Framework Programme for R&D (FP7). This document does not represent the opinion of the European Union, and the European Union is not responsible for any use that might be made of its content.
We would like to thank all teachers and students participating in our workshops.

References

[1] Brooke, J. SUS-A quick and dirty usability scale. Usability evaluation in industry, 189(194), 1996. pp. 4-7.

[2] de Jong, T., Sotiriou, S. and Gillet D. Innovations in STEM education: the Go-Lab federation of online labs. Smart Learning Environments, 1(1), 2014. pp. 1-16.

[3] DePaula, R. Lost in translation: a critical analysis of actors, artifacts, agendas, and arenas in participatory design. In Proceedings of the eighth conference on Participatory design: Artful integration: interweaving media, materials and practices-Volume 1, July. ACM, 2004, pp. 162-172.

[4] Finstad, K. The Usability Metric for User Experience. Interacting with Computers, 22, 5, September 2010.

[5] Govaerts, S., Cao, Y., Vozniuk, A., Holzer, A. C., Garbi Zutin, D., San Cristóbal Ruiz, E., Bollen, L., Manske, S., Faltin, N. and Salzmann, C. Towards an Online Lab Portal for Inquiry-based STEM Learning at School. Proc. of ICWL 2013, 2013.

[6] Hassenzahl, M. and Monk, A. The Interference of Perceived Usability From Beauty, Human Computer Interaction, 25:3, 2010, pp. 235-260

[7] Heintz, M., Law, E. L.-C., Govaerts, S., Holzer, A. and Gillet, D. Pdot: Participatory Design Online Tool, CHI '14 Extended Abstracts on Human Factors in Computing Systems. ACM, 2014, pp. 2581-2586.

[8] Heintz, M., Law, E. L.-C. and Heintz, S. Review of Online Tools for Asynchronous Distribute Online Participatory Design. In Dominik Renzel, Ralf Klamma (eds.), "Large-Scale Social Requirements Engineering", IEEE Computer Society Special Technical Community on Social Networking E-Letter, vol. 2, no. 3, September 2014.

[9] Nielsen, J. Usability engineering. Elsevier, 1994.

[10] Ovaska, P., Rossi, M. and Smolander, K. Filtering, negotiating and shifting in the understanding of information system requirements. Scandinavian Journal of Information Systems, 17(1), 7. 2005.

[11] Schuler, D. and Namioka, A. Participatory design: principles and practices. 1993.

Guidelines to Specify HCD Activities in the Call for Tender for Public Administration Websites

Simon Mastrangelo[1], Rosa Lanzilotti[2],
Maurizio Boscarol[3], Carmelo Ardito[2]
[1]Ergoproject - Studi ergonomici applicati Srl
Via Antonio Pacinotti 73/B, 00146, Roma
[2] Dipartimento di Informatica, Università degli Studi di Bari Aldo Moro
Via Orabona, 4, 70125 Bari, Italy
[3] www.usabile.it
s.mastrangelo@ergoproject.it,
rosa.lanzilotti@uniba.it,
maurizio.boscarol@gmail.com, carmelo.ardito@uniba.it

Abstract. Despite studies that document the economic benefits of usability evaluation beyond the obvious improvement in user satisfaction, this research in human-computer interaction has too little impact on software development practice. Consequently, many software systems developed show a very poor usability. This is true equally for commercial software and for public administration web sites. Government websites have the objective of improving the interaction between citizens and public agencies or government functions. And yet, our study of software engineers revealed an important but previously hidden result, namely: during product development process companies focus almost exclusively on the requirements formally established in the Call for Tender (CfT) of a particular project. Generally, such requirements do not include usability and user experience (UX) and, thus, those qualities are neither designed into the project, nor evaluated on the way to completion and delivery. In this position paper, we review the motivation for including UX and usability requirements in CfTs, suggest some beginning steps for how to specify these qualities in the CfTs and show the initial effort for how to evaluate the efficacy of these specifications and guidelines.

1 Introduction and motivation

The Digital Agenda, presented by the European Commission, is one of the seven pillars of the Europe 2020 Strategy. It defines the objectives for the growth of the European Union (EU) by 2020. Specifically, it proposes to better exploit the potential of Information and Communication Technologies in order to foster innovation, economic growth and progress.

A composite index, called Digital Economy and Society Index (DESI), summarizes relevant indicators on Europe's digital performance and tracks the evolution of EU member states in digital competitiveness [9]. The DESI

includes 5 main dimensions: 1) *Connectivity*, which measures the deployment of broadband infrastructure and its quality; 2) *Human Capital*, which measures the skills needed to take advantage of the possibilities offered by a digital society; 3) *Use of Internet*, which accounts for the variety of activities performed by citizens already online; 4) *Integration of Digital Technology*, which measures the digitization of businesses and their exploitation of the online sales channel; 5) *Digital Public Services*, which measures the digitization of public services, and focuses in particular on e-Government and e-Health.

Figure 1 shows the score of each DESI dimension related to 2014 and 2015. In 2015, the overall DESI score is 0.48, which highlights an improvement in digital development in comparison to the previous year, when it was 0.45: this shows that both the EU as a whole as well as individual member states are progressing towards a digital economy and society.

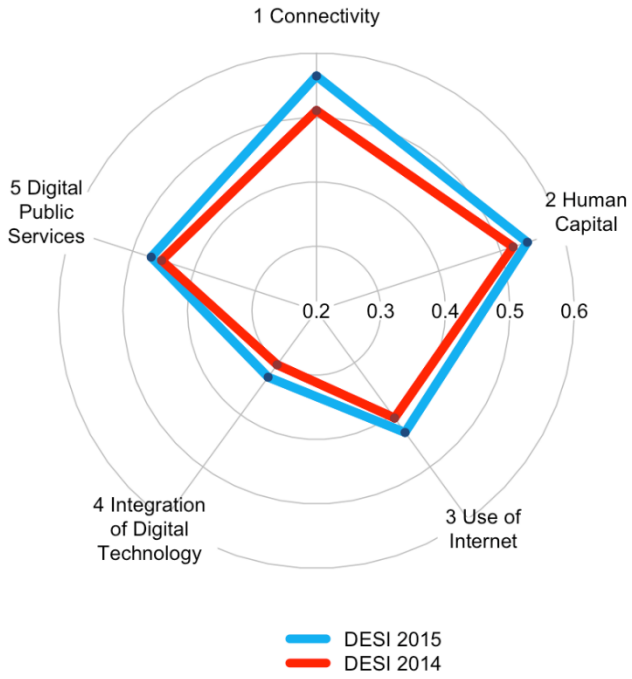


Figure 1. The score of the DESI dimensions related to years 2014 and 2015 [from https://ec.europa.eu/digital-agenda].

However, member states are at different levels of development and are progressing at different speeds. Digital Public Services is the dimension where the states' performance is most fragmented. Denmark and the Netherlands are among the leaders in the online public services not only in Europe but also in

the world, in contrast to Italy and Slovenia, countries that are moving more slowly toward digital public services.

The Italian level of use of digital public services is still under the European average: only 18% of Internet users interact with Public Administration (PA) websites. This puts Italy in twenty-fifth place among the 28 EU member states measured by the DESI. PA websites are judged poorly because of incomplete implementation of digital services, and poor usability [6].

In order to improve this situation, every PA website should be designed for user experiences (UX) that enable citizens to assert their rights and fulfill their duties, and evaluated for usability. Usable websites that provide a positive UX to their users is both a formal and ethical responsibility. Italy has had legislation and policies that require PA to work toward usable and accessible services since 2001 [4]. Transparency in government is a goal that can be accomplished for e-Government and Open Government. If embraced fully, enormous amounts of information/services could become available to every citizen, regardless of his/her status, employment, age, education level and on any device he/she uses (e.g., computer, tablet, smartphone). How shall we provide motivation to software developers or organizations hired to produce PAs that Human-Centred Design (HCD) methods produce usable websites.

In the last few years, a considerable part of our research has been devoted to this aim. Our study [1] called out an important problem about the reasons companies do not address usability engineering during their product development process: many companies focus almost exclusively on the requirements formally established in the project Call for Tender (CfT) (comparable to a Request for Proposal RFP). Generally, such requirements do not include usability and UX and, thus, those methods and activities are not built into the plans, the customer acceptance criteria, nor as part of the system requirements. What's missing are guidelines about User Experience and usability to help the PA to introduce HCD methods in their CfTs in order to improve measures of usability and usefulness for websites or other digital public services.

Since July 2013, we have been collaborating with members of the GLU (Gruppo di Lavoro per l'Usabilità), an Italian working group on usability related to the Italian Ministry of Public Administration, in order to improve the usability of PA websites and other e-government systems. In May 2015, GLU published a new version of a document that provides detailed guidelines for the design of the websites of the PA, called eGLU Protocol 2.1 [8], whose aim is to guide web masters and web editors, who do not have experience on usability and UX evaluation, in the identification of usability problems of the

websites they work on, by committing limited resources in terms of time and people.

This is very important, but it is not enough. Government website designers should make usability a priority [5]. In according to [5], creating usable website means: 1) making the experience familiar: sites have to conform to evolving web design norms; 2) understanding citizens: websites should be organized around citizens' needs; 3) practicing appropriate consistency: citizens have to feel they were in an integrated place when they left home base; 4) testing and evaluating website designs: usability can be reached only if the website prototypes designs are evaluated from the early stages of the software lifecycle.

Different countries have defined guidelines aiming at supporting designers in creating usable government websites. See, for example, guidelines proposed by United States of America (see [10]), but also those one issued by Italy [4]. In this position paper, we illustrate a set of guidelines to introduce the HCD act in the CfT for the PA websites.

The remainder of this paper is structured as follows. Section 2 reviews our analysis of Italian and international CfTs to determine whether and how HCD, UX and usability are mentioned. Section 3 presents the guidelines we have developed to integrate HCD software development practices in CfTs. Finally, Section 4 describes an ongoing experiment in including HCD in a contracted software project for digital public services.

2 HCD in the Call for Tenders

A considerable part of our research in the last years has been devoted to methods for addressing usability and UX in software development, trying to convince software practitioners of the advantages of integrating HCD techniques in their development practices. To this aim, we conducted an experimental study with software companies to investigate the use of HCD in their software development processes and, in particular, to analyse how they address usability and UX of the products they create [1].

As mentioned above, software companies attend to requirements mentioned in the CfT, and put aside any other concerns they might include in a differently drawn project. To understand how often UX and usability are mentioned in CfTs, we have performed an analysis of 44 CfTs for ICT systems issued by public and private organizations (26 in Italy and 18 in International countries) in order to verify to which extent the CfTs explicitly requires the use of HCD techniques. Table 1 shows the preliminary results revealing 3 different categories.

		Category		
	Organization	**Specific requirements**	**General Requirements**	**No requirements**
Italian CfT	*Public*	0	12	4
	Private	1	9	0
International CfT	*Public*	2	7	8
	Private	0	1	0
Total		**3**	**29**	**12**

Table 1. Preliminary results of the analysis investigating HCD requirements in CfTs.

Only in 3 CfTs analysed (i.e. 1 Italian CfT issued by one private organization and 2 CfTs issued by international private companies), formal usability tests are specifically required (see Specific requirements in Table 1). For example, in the CfT issued by the International Centre for Trade and Sustainable Development (ICTSD) for the redesign of their website [3] it is written: "*Usability: Clarity of interaction between the users and the site is crucial, and must be paramount in the redesign. The web team will run at least two usability tests throughout the design and development process of the new site (to create a short list of the most serious problems and a commitment to fixing them before the next round of testing). A specific usability testing plan (with details on number of participants in each round, who we test with, where/when we test, who watches, reporting, etc.) will be released before the beginning of the design process*." The CfT also stated that the most serious problems identified in the first test have to be fixed before running the second usability test.

Twenty-nine analysed CfTs (i.e. 12 CfTs issued by Italian public institutions; 9 CfTs by private Italian companies, 7 CfTs by International Public institution and finally 1 CfT issued by International private institution) refer to HCD techniques as a general requirement. For example, in the CfT of an Italian PA for the development of a system for registering employees presence there is only a sentence that refers to usability [2]: "*Application programs should preferably comply with the quality requirements, as those reported in the ISO 9126 (i.e, functionality, reliability, usability, efficiency, maintainability and portability)*". The word "preferably" indicates that software quality is not mandatory. In particular no specific requirements about usability are provided, so it is not clear if and how companies would attempt to fulfill or evaluate this aspect.

Finally, 12 of the analysed CfTs (i.e. 4 CfTs issued by Italian public institutions and 8 CfTs by International public institution) did not mention any requirement related to HCD, usability or UX.

From these several studies, there is little or no attention to HCD, usability and UX issues in the CfTs, since in most cases relevant requirements are poorly specified, or are not valid and verifiable. Without specific requirements for UX or usability in the CfT, the software companies will not use HCD techniques in their projects.

3 Guidelines to Introduce HCD in the CfTs

HCD activities and techniques can be specified in a CfT at different phases of evaluation and selection of a contractor [7]:

- *Qualification phase* - it requires tenderers to demonstrate previous experience and/or expertise in HCD (e.g., the company has employees that are expert in usability, or company has evaluated their own products previously).
- *Selection phase* - a bonus is assigned (in terms of points during evaluation of response to a CfT) to companies whose software development lifecycle is based on an HCD approach.
- *System requirement phase* - HCD specific requirements are included within the technical specification documents that the contractor must fulfill. Even more specific would be usability measures as part of a customer acceptance test.

Each of the three phases increases the quality of the CfTs and of the activities contractors have to carry out. Table I shows a series of more complex or complete HCD activities the contractor has to perform: User requirements analysis, Usability and/or UX monitoring, Navigational scheme test, and Usability/UX test. We've proposed three different levels of requirements related to website complexity:

1. *Basic requirements* – to be included in the CfTs related to websites of low complexity with reduced budget and execution time; they require preliminary evaluations with some qualitative usability assessments.
2. *Medium requirements* – to be considered in the CfTs related to websites of medium complexity and medium/high budget and execution time; they require to specifically identify the target users, monitor the perceived usability (through questionnaires) and navigational scheme along with some qualitative assessments of usability.
3. *Advanced requirements* – to be included in the CfT related to websites from medium/high to high complexity and medium/high budget and

execution time; they are similar to the Medium requirements but with more attention to quantitative as well as qualitative assessments.

For example, if the requirement level is "Basic", the User requirement analysis is performed using questionnaires and interviews in order to gather the most important characteristics of the target users, their needs, etc. While if the level is "Advanced" a specific user profiling is required through techniques, such as personas and scenarios.

Activity	**Basic requirements**	**Medium requirements**	**Advanced requirements**
User requirements analysis	Questionnaires Interviews	Personas Scenarios	Personas Scenarios
Usability and/or UX monitoring		Online questionnaire for monitoring usability	Online questionnaire for monitoring UX
Navigational scheme test		Card sorting or reverse card-sorting	Card sorting or reverse card-sorting
Usability/UX test	At least 2 formative test involving 5 users for 6 tasks Questionnaires for assessing the perceived UX	At least 2 formative test involving 5 users for 8 tasks Questionnaires for assessing the perceived UX	At least 2 formative test involving 5 users for 8 tasks Questionnaires for assessing the perceived UX Summative test involving at least 15 users

Table 2. The three levels of HCD requirements.

In the following, there is an example of a section, which can be inserted in the CfT related to PA websites, illustrating the basic requirements related to the HCD activities.

Human-Centred Design, usability and User Experience

In order to follow the indication provided by the 2011 Guidelines for the PA Websites (Chapter 4.4.2 "Usability")[1],

[1] Downloadable from http://www.funzionepubblica.gov.it/lazione-del-ministro/linee-guida-siti-web-pa/indice/cap4-garantire-la-qualita-dei-siti/indirizzi-sullusabilita.aspx

the contractor must specify the methodology and the process to follow in the evaluation and implementation of the website (ISO 9241-11) through the use of the Human-Centred Design approach (ISO 9241-210).

This process considers at least the following activities:

1. Identification of the specific characteristics of target users through the involvement of their representatives in the definition of requirements through interviews and/or questionnaires
2. Execution of at least 2 formative usability tests, with a minimum of 5 participants and 6 tasks for each test to be carried out during the development of the prototypes, wireframes or pre-release versions of the website, in order to identify the main usability problems and to have them fixed before the final release. The types of participants and the navigation tasks to be used during the test must be proposed by the contractor and approved by the customer. Participants involved in the second test should be different from those involved in the first. The results must be documented by a report including:
   - Number of participants, their demographic characteristics, and the navigation tasks they performed;
   - Success index;
   - List of the identified problems (with possible solutions) and their seriousness;
   - Subjective metrics (es. SUS, Umux).

The execution of these activities is compulsory to consider the implementation of the work complete, but there is no restrictive covenant respect to the level of usability measured.

For clarification of the used terminology in this section, please refer to the Section "Terms and definitions" of following technical standards:

- ISO 9241-11 " Ergonomics of human-system interaction - Guidance on usability"
- UNI EN ISO 9241-210 - Ergonomics of human-system interaction -- Part 210: Human-centred design for interactive systems
- ISO/TR 16982 - Usability methods supporting human-centred design
- ISO/IEC 25062 - Common Industry Format (CIF) for usability test reports

4 Current work and anticipated next steps

This paper has presented a proposal for guidelines to specify the execution of HCD activities by contractors in the calls for tender related to Public

Administration websites. The guidelines are based on three different levels of HCD requirements that are characterized by the growing website complexity. A first application of the advanced requirements is currently ongoing. Specifically, these guidelines were included in the CfT for the redesign of a government platform. A restricted procedure involving 5 selected software companies was carried out. The contracting company was glad to introduce HCD techniques in the platform development process. The company did not propose HCD techniques in its proposal, because the cost associated would have priced them out of consideration for the contract. The company has already defined personas and scenarios for the user requirement analysis activities and now platform prototypes have been created. At the workshop, other results of the public platform development developed according to HCD approach will be presented.

It is worth highlighting that, at this stage, our aim is to test/share the proposed guidelines both with the PA (for case studies) and with professionals/researchers to improve and complete them. The execution of the prescribed HCD or usability techniques is compulsory to consider the implementation of the work complete; however, possible usability problems detected will not prevent the final acceptance of the system.

Acknowledgements

This work is partially supported by the Italian Ministry of University and Research (MIUR) under grant PON 02_00563_3470993 "VINCENTE", by the Italian Ministry of Economic Development (MISE) under grant PON Industria 2015 MI01_00294 "LOGIN", and by PON04a2_B "EDOC@WORK3.0".

We would like to thank also the Emilio Simonetti of the Dipartimento della Funzione Pubblica (Italy) and Alessandra Cornero of the FormezPA (Italy). A special thank to the reviewers for their valuable comments and to Nancy Frishberg for her careful proofreading and precious suggestions.

References

[1] Ardito C., Buono P., Caivano D., Costabile M. F. Lanzilotti R. Investigating and promoting UX practice in industry: an experimental study. International Journal of Human Computer Studies 72, 6, 542-551, 2014.

[2] Call for tender - Sistema di rilevazione presenze del personale della Regione Puglia. http://www.regione.puglia.it/index.php?page=burp&opz=getfile&file=17.htm&anno=xl&num=166. Last access: June 4th, 2015

[3] Call for tender - Redesign the Corporate Website. http://ictsd.org/redesign/wp-content/uploads/2013/09/i-Call-for-Tenders_Design.pdf. Last access: June 4th, 2015
[4] Circolare del Dipartimento della Funzione Pubblica, 13 marzo 2001, n. 3, "Linee guida per l'organizzazione, l'usabilità e l'accessibilità dei siti Web delle pubbliche amministrazioni"; Decreto Legislativo 7 marzo 2005 n. 82, Art. 53; Linee Guida per i siti web delle PA, 2011, capitolo 4.4 "Accessibilità e usabilità"; Decreto Legislativo 14 marzo 2013, n. 33, Riordino della disciplina riguardante gli obblighi di pubblicità, trasparenza e diffusione di informazioni da parte delle pubbliche amministrazioni Art. 6 "Qualità delle informazioni".
[5] Gill, K. Usability of Public Web Sites. In Usability in Government Systems, edited by Buie E. and Murray D. Elsevier, USA, 2012, pp. 21-40.
[6] https://ec.europa.eu/digital-agenda/en/scoreboard/italy. Last access: June 4th, 2015
[7] Jokela, T. and Buie, E. Getting UX into the Contract. In Usability in Government Systems, edited by Buie E. and Murray D. Elsevier, USA, 2012, pp. 251-265.
[8] Protocollo eGLU. Available at: http://www.funzionepubblica.gov.it/glu.aspx. Last access: June 4th, 2015
[9] The DESI index. Available at: https://ec.europa.eu/digital-agenda/en/digital-economy-and-society-index-desi
[10] U.S. Dept. of Health and Human Services. The Research-Based Web Design & Usability Guidelines, Enlarged/Expanded edition. Washington: U.S. Government Printing Office, 2006.

Design Requirements for Web Applications to Affect the End User Emotional State

Giulio Mori, Fabio Paternò, Ferdinando Furci
ISTI – CNR
Via Moruzzi 1, 56126, Pisa, Italy
{giulio.mori, fabio.paterno, ferdinando.furci}@isti.cnr.it

Abstract. We report on work aiming to identify Web design requirements to stimulate specific emotions. The results are based on a user test analysis with 50 participants evaluating the emotional impact of six different Web design criteria eliciting respectively hate, anxiety, boredom, fun, serenity, love. The design criteria applied come from a previous survey with 57 participants suggesting the main typical emotions during Web interaction and associating them with some specific Web characteristics. An additional user test with 40 users investigated deeper which design features are most important to elicit a particular emotional state. A formalization of such criteria can be important to support Web designers developing Web applications taking into account the user's emotional state.

1 Introduction

The role of emotions has acquired more and more importance in HCI [1, 2]. Emotions are complex and depend on individual preferences, attitudes, moods, affect dispositions, and interpersonal stances; "there is no single standard gold-method for their measurement" [3]. Unfortunately, even if in HCI there is a general agreement about the importance of the user emotional state during the interaction, no concrete formalization of the relevant design criteria has been proposed in order to support designers aiming to obtain applications able to stimulate the user emotional state. There are examples of emotions classifications, such as Geneva [4], or contributions comparing different versions of Web pages [5] to investigate the impact of their attractiveness [6] or aesthetics, but no work has focused on typical emotions during Web interaction. Our goal is to investigate the impact of some Web design criteria to elicit a particular emotional state on the user and if some specific Web design features are more effective than others. To accomplish this goal, we conducted one survey and two user studies. The preliminary survey aimed to collect same basic indications from a sample of 57 users. In particular the survey intended to understand the most recurring emotions during the Web interaction and the associated Web design features for each proposed emotion.

The final data produced a scale of six emotions (ordered from the perceived most negative to the most positive) typical of Web interaction: hate, anxiety, boredom, fun, serenity, love. After the survey, we designed and implemented six interfaces to which we applied the collected criteria, and we asked 50 users to test the emotional effectiveness of the Web design criteria applied. Finally, in order to understand if some Web aspects of design can be more relevant than others to stimulate a specific emotional state, we performed another user test with 40 users. Section 2 analyses in details the survey and the two user tests, and provides an overview about the future evolution of this research, and then the conclusions are drawn.

2 Web Design Criteria able to Stimulate Emotions

The process for identifying some Web design criteria able to stimulate an emotional reaction on the user, has been defined through many steps.

2.1 A Survey: Collecting Opinions about Emotional Web Design

The 57 participants (25 females and 32 males) had an average age of 38,21 years (ranging from 26 to 59). The sample considered users with heterogeneous educational level and Web development experience. We asked each participant to suggest the typical emotions (maximum 8) involved during Web interaction. The only constrain for each proposed emotion was to indicate also the opposite one (depending on the negative or positive emotional valence, in their perception). The average number of proposed emotions was 3.84 per user. On the base of their perception, users had to order each proposed emotion, and to associate each one with specific characteristics in terms of main Web design aspects (such as colours, page structures, contents distribution, type of media, navigation & interaction elements, etc.). We collected 219 emotions perceived as negative and 219 perceived as positive.

2.1.1 Results

Analysing the proposed emotions, we discarded synonyms, emotions with a low number of preferences or having the same Web characteristics suggested by different users. The resulting six emotions may not be exhaustive. However, the goal of this work is not to provide a further emotion classification, but rather if possible, to identify some clearly distinguishable design characteristics that could elicit a specific emotion. As a final result, we obtained a filtered scale of six emotions: hate, anxiety, boredom, fun, serenity,

love. Hate and love express the sense of disliking/liking for something (e.g. typical of Web social networks). Anxiety and serenity express the emotional state during critical/safe actions (i.e.: the user is booking/buying something on Web inserting personal or credit card data). Boredom or fun depend on the way the contents are presented attractively. For each emotion, the users suggested us some Web design features as potentially able to stimulate the considered emotional state [7]. Even if the proposed Web design features are distinguishable, emotions close in the scale have some common features. This indicates the complexity of the emotions and the fact that there is not a well definite distinction between similar (but not equal) emotional states.

2.2 A User Study: Testing the Emotional Impact of the Criteria

After having performed the preliminary survey, we wanted to check concretely if the collected data were effectively efficient to elicit each considered emotion. To accomplish this goal, we implemented six Web interfaces applying the Web design criteria suggested in the survey. Each Web interface presented the same content (except very minimal additions suggested by the users) in a different design style. The goal of each interface was to stimulate one of the six emotions. Fifty new users (21 females an 29 males) evaluated the emotional impact of the six interfaces through a user test. The average age was 38,28 years (ranging from 26 to 77). The sample considered users with heterogeneous educational level and Web development experience. We chose the Beatles' musical history as topic for the six interfaces. The interfaces contained a short textual biography, a player where the user could listen to five famous songs, a video, a form to buy virtual tickets for revival events and six clickable graphic covers of famous albums. The 50 users evaluated in random order each one of the six interfaces giving a judgement in a scale from 1 to 5 scale (where the value 1 indicated that the page was very ineffective to elicit the target emotion, while value 5 indicated that the page was very effective, and the value 3 represented the neutrality). Before giving their judgments, users had to perform three tasks for each interface: a) find the answer to a question in the biography (so the user could test the reading) b) clicking one of the six albums (so the user could test the navigation elements) c) filling in the form (so the user could test the interactive elements). The six Web interfaces had been designed as follow: a) the interface eliciting hate had the content distributed in one single page, a confused layout, and blurred text and elements; b) the interface eliciting anxiety presented the contents distributed in multiple pages, showed intermittent light effects and jerky

transformations, with a countdown as a pressure factor to fill in the form; c) the interface eliciting boredom was one single page with long text, without images or videos, requiring more fields to fill in the form; d) the interface eliciting fun had the contents distributed in multiple pages, and showed unpredictable animations and dynamic effects; e) the interface eliciting serenity was very simple to minimize the user's effort, allowing easy navigation through TABS elements; f) the interface eliciting love was a long page with an appealing graphics and was usable [7, 8].

2.2.1 Results

The user appreciated positively the emotional impact of the interfaces. In particular, the average judgment and the deviation standard (SD) related to the six designs were: 4.48 (SD 0.64) for hate, 4.52 (SD 0.61) for anxiety, 4.16 (SD 0.88) for boredom, 3.32 (SD 1.21) for fun, 4.22 (SD 0.72) for serenity and 3.64 (SD 1.06) for love. Fun and love were considered the more critical Web design criteria, even if the judgment was on the positive side. In particular, the presence of animations/dynamics effects applied for fun was considered excessive, and the red and pink colours applied for love were considered too shocking.

2.3 A User Study: Looking for Relevant Emotional Features

Encouraged by the positive results of the previous user test, we wanted to understand if some Web design features were more relevant to stimulate an emotional reaction. For this goal, we designed and implemented a Web application able to adapt its design to the emotional state selected by the user. For each emotion we designed two versions (implemented with different Cascading Style Sheets): the first version applied the criteria of the previous user test but improved by the users' suggestions; the second version presented small differences for those elements deemed controversial in the previous user test (such as colours, visual characteristics, blurred or clear text and dimension of images/videos). We recruited other 40 users and we divided them in two groups A e B of 20 users each. Group A (11 females and 9 males) evaluated first version for each emotion and Group B (14 females and 6 males) the second one. Both groups were heterogeneous in terms of education and developing experience. Each user evaluated one version of each design in random order. For each emotion, every user had to perform three similar tasks of the previous user test, and had to give a judgement (in a scale from 1 to 5) about the emotional impact, choosing also the main three aspects of Web

design (from a proposed list) considered fundamental to stimulate the target emotional state.

2.3.1 Results

The results of the two groups appeared consistent [8]. The judgments on the emotional impact improved with respect to the results of the previous user test. Only the boredom evaluation decreased due to the fact that even if the biography was long, the actual information to find was at the beginning. The main aspects for the six Web designs resulted as follow: a) *hate*: **confused layout, difficult interaction and navigation**; b) *anxiety*: **stress factors, blurred text/images/videos**, and **dynamic effects**; c) *boredom*: **excessive information, absence of dynamic effects** and **absence of images/videos**; d) *fun*: **appealing graphics/aesthetics, dynamic effects** and **colour images/videos**; e) *serenity*: **ordered layout, reassuring elements** and **easy interaction & navigation**; f) *love*: **appealing graphics/aesthetics** and **the reassuring elements**.

2.4 Discussion

Our user tests have showed that usability, even if it is an important factor, it is not the unique aspect responsible to elicit an emotional reaction on the user (e.g. different interfaces with poor usability, can produce different emotional reaction, as different interfaces with good usability can produce others). Users considered positive emotions important to improve the user experience, but in some specific application domain also negative emotions have been considered useful (e.g. in educational, learning or psychological tools where it is necessary to increase the awareness of children about the difference between good and bad behaviours).

2.5 Possible Approaches to Emotion Changes

The final goal of this research aims to obtain solutions able to adapt the design of Web applications to stimulate more positive emotions. Thus, we need to understand the best approach to support a transformation of design criteria for this purpose. The idea of a transformation from a design to another has been suggested by the users in the preliminary survey as a gradual process. Choices of strategies (which elements of the interface can be transformed before others and the evaluation of the times of the transformation) are important to avoid the perception of a too traumatic change of Web design. Depending on the original and target emotional design, the common elements between two different designs can facilitate a smooth perception of transformation on the

user, otherwise an intermediate design could be necessary. We plan new user tests to understand more concrete indications about more effective strategies to guide the user towards a different emotional state.

3 Conclusions and Future Work

The process for identifying some Web design criteria able to stimulate emotional reactions during the interaction has produced encouraging results, showing that different Web design requirements are able to stimulate different emotions. However, further investigation is necessary. The ultimate goal of this research is the formalization of a set of design criteria for Web interfaces able to stimulate effectively emotional states. In addition, we plan to monitor the emotional state of the user and the level of attention (with some physiological sensors, an eye tracker and log analysis) to design Web applications able to adapt the design in order to guide the user towards a more positive emotional state and a pleasant user experience.

References

[1] Karlsson, M.: Expressions, Emotions, and website design. CoDesign, 3, 1 (2007), 75-89.
[2] Kim, J. Lee, J., Choi, D.: Designing Emotionally Evocative Homepages: An Empirical Study of the Quantitative Relations Between Design Factors and Emotional Dimensions. International Journal of Human-Computer Studies, 56, 6 (2003), 899-940.
[3] Scherer, K. R.: What are emotions? And how can they be measured? In: Social Science Information, 44, 4 (2005), 695-729.
[4] Bànziger T., Tran V., Scherer K.R.: The Geneva emotion wheel, Journal, Social Science Information, 44, 4 (2005) , 23-34.
[5] Kim, J. Lee, J., Choi, D.: Designing Emotionally Evocative Homepages: An Empirical Study of the Quantitative Relations Between Design Factors and Emotional Dimensions. International Journal of Human-Computer Studies, 56, 6 (2003), 899-940.
[6] Hartmann, J., Sutcliffe, A., De Angeli, A.: Investigating Attractiveness in Web User Interfaces. In: Proc. CHI 2007, ACM Press (2007).
[7] Mori, G., Paternò, F., Furci, F.: Design Criteria for Web Applications Adapted to Emotions. In: ICWE 2014. Springer LNCS Vol. 8541, 400-409, 2014.
[8] Mori, G., Paternò, F., Furci, F.: Design Criteria for Stimulating Emotions in Web Applications. In: INTERACT 2015.

From Usability Workarounds to Usability Around Work

Kimmo Tarkkanen, Ville Harkke
Information Systems Science
Turku School of Economics, University of Turku
Rehtorinpellonkatu 3, 20500 Turku, Finland
kimmo.tarkkanen@utu.fi, ville.harkke@utu.fi

Abstract. In this paper, we introduce usability workarounds that take place in system design during IT tendering, and usability evaluation around work that characterizes an iterative IT product development. We claim that highly formal and detailed verification criteria and procedures in tendering cause usability workarounds by the IT vendors. To avoid workarounds and to bring usability activities more around work during IT tendering, we suggest applying open-ended usability test tasks for system validation and usability problem qualifiers as criteria.

1 Introduction

Many public information system purchases are subject to tender. Request-for-proposals (RFP) are documents that define a set of desired system requirements and a selection criterion for the proposed systems. RFP creation is a critical phase of governmental IS procurement also in terms of usability [1]. Jokela et al. [1] suggests that only user performance-based usability requirements can be verifiable, valid and comprehensive enough for RFPs. A performance requirement is e.g. how quickly the user should be able to accomplish a certain task. User performance can be measured in terms of effectiveness, efficiency and satisfaction i.e. the elements of the definition of usability [2].

We maintain, however, that applying very detailed performance requirements in the RFP may lead to distorted usability design at the IT vendor's site. HCI research has concentrated on the creation of RFP from the IT purchasing organizations' point of view, but how the usability practices of the vendors emerge and transform for tendering purposes are less studied. Here we present problems of usability design practices in a large and complex IT tendering case and then seek possible solutions and new openings from an agile, user-centred IT development case.

2 Two types of usability work

2.1 Usability workarounds

We participated in an IT tendering case as usability evaluation designers and researchers at the vendor's site [3,4]. We were responsible for measuring and improving the usability of the proposed IT system before it was introduced to the purchasing organization and exposed to their usability evaluation. The RFP contained formal user performance and satisfaction requirements and metrics as well as detailed information about how these were to be evaluated (e.g. usability tests tasks were published to the vendors). Thus, this allowed us to copy the evaluation practices (phase 1 in Figure 1) and prepare the system (phases 2-3) to best correspond with the requirements (phases 4-5). The usability criteria in the RFP included effectiveness as task completion rate (percentage), efficiency as number of errors made, time on task and number of interaction steps used, and satisfaction of users as measured with questionnaires (phase 0). These variables had nearly equal weights in overall system selection. A usability requirement was represented as, for example, a test task given to users to "*send an email message to John Smith*" (where the system could be a webmail). The buyer then evaluates how well their experienced email users perform the task with different webmail systems in usability tests (i.e. a common evaluation situation in system purchases for expert domains; phase 6).

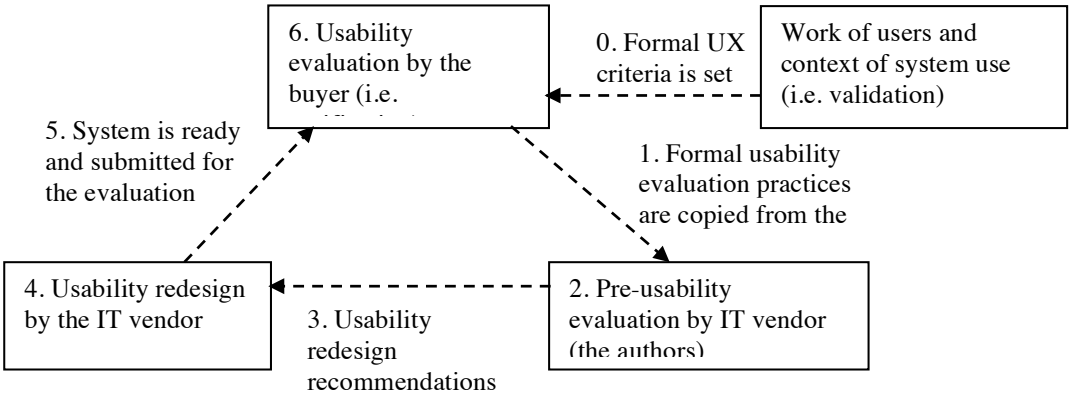


Figure 1. Usability evaluation setting in the tendering case.

The motivation and goal of vendors' usability work is success in tendering. The concept "workaround" presupposes that there is some sort of obstacle in achieving this goal that must be circumvented and hereby the 'standard

operating procedure' is bypassed or overridden, i.e. the work is carried out some other way than usually [5]. Thus, a usability workaround means bypassing human-centred design principles and usability guidelines and carrying out usability work some other way than usually due to confronted obstacles.

What are the obstacles the IT vendor may confront? First, usability design is limited by the time given and restrictions to modify the system. Second, the communication limitations with the buyer hinder usability design, because the vendor may want to know e.g. should John Smith be able to receive the message or who decides when the sending is completed (evaluators when send-button is pressed or when users verbally express?). Is the task completed successfully when a message is sent without content or title? These questions must have puzzled the buyer also, because usability requirements, procedures and criteria in the RFP were very detailed and rigid. Table 1 presents how certain characteristics of the RFP were confronted as obstacles by usability designers, implying usability workarounds and some further problems.

Table 1. Vendor's usability workarounds as responses to the RFP characteristics.

Characteristics of the RFP	**Usability workarounds**	**Problems**
Tendering schedule is tight	Perform only one evaluation-design iteration	Only minor and superficial changes can be implemented.
Usability test tasks are published	Design system usability to correspond the tasks	Task design should avoid giving clues to task completion.
Communication is limited	Interpret requirements to your system's advantage (when most efficient & effective)	If test tasks are subject to many interpretations, this may raise conflicts after the tendering.
Usability test tasks are low-level and strictly sequenced	No special workarounds. Design your system to support the requested task sequence.	The system may be more efficient and effective in another task sequence than requested.
Usability test tasks have interconnections and elements that recur	Analyse which test tasks are the most critical for other tasks to be completed efficiently and put design effort on the critical ones.	Design effort is misdirected from criticality in the real work.
Effectiveness is measured as task completion rate	Modify the terminology of the system (e.g. menu items, function names) and information appearance (dropdown lists) to correspond with the test tasks. Set also the critical views to be visible/open by default.	Renamed features, data ordering and visibility settings may serve only the test task, yet be inappropriate in other work tasks.

Efficiency is measured as execution time or as the number of navigation errors	Add shortcuts to shorten the navigation path despite of users getting lost in the system. Especially applicable when the next task begins from the main view. Inactivate features that are on the unrequested navigation path.	Users should be in control and informed about the system state and place (e.g. Nielsen's heuristics [6]).
User satisfaction is measured	Avoid showing error messages (e.g. by inactivation) not to give a feeling of erroneous system	In complex systems, error dialogs are inherent part of the systems.

The examples of different usability workarounds to improve task completion and performance efficiency were solely based on the RFP. Although such usability workarounds are possible for the system demonstration and testing purposes the resulting system configuration may not be feasible during the real use after the tendering. Therefore, formal UX definitions of the RFP must be based on an in-depth understanding about the current and desired work practices and needs of the users (i.e. validation). Definitely the buyer organization in this case possessed that knowledge. The main problem was the low-level and very detailed formulations of test tasks (i.e. descriptions users' work) and rigid requirements to comply with these. An implication was that the IT vendor focus was on usability verification (with usability workarounds) instead of on usability validation as a fit for purpose.

2.2 Usability around work

On the other extreme of the usability work, we partook in an agile development project of a mobile electronic patient record application for nurses in hospitals wards. This development was not yet subject to tender, but a traditional R&D activity of an IT company where such usability workarounds are not needed. Compared with the tendering situation, here was a closer user-developer relation and direct access to users during the usability work.

We conducted the first think-aloud usability tests of the paper prototype (although as pictures in a mobile tablet platform) in-the-wild with four nurses. Tests were preceded by an extensive user research by the application developers. Yet, and in contrast to the tendering case, our test approach was highly open-ended: The test task given to users was to '*arrive at workplace and begin to prepare your work shift with the [prototype] as a new application available to use*'. However, test administrators could pose additional questions ex tempore during the session. The motivation for the open-ended task was that a) users would express their needs more freely [7] and b) test would be based on real (yet simulated) tasks and, c) by not-restricting the view on the

application only (but covering the users' own articulation of their work) the test would challenge and help the continuous requirements elicitation process of the agile development.

Later, we categorized these findings based on usability problem qualifiers (i.e. failure qualifiers by [8]) (see also [9]). The "qualifier" attribute helps designers to see the real problem (i.e. what is wrong with the initial design) [8]. Ten categories were identified, which here point out that a single usability intervention – when it is practiced around work and freely in the collaborative environment unlike in the previous case – provides a broad spectrum of findings for usability verification and validation. Moreover, we want to draw attention to possible application of categories in tendering situation where the classification of usability problems could serve as a basis to set criteria in the RFPs.

Four out of ten categories represent poor fit in system-work relation: System functionalities or information elements are *missing*, *unnecessary*, *inadequate*, or *misplaced*. These are rather common findings from usability tests, yet for example missing functionalities and information items are hard to identify with classical, very low-level and only "doable" test tasks. Further categories include *misinterpreted* which represents problems in user-system relation i.e. users misinterpret the system (a single element of it or the whole purpose of it), *open design issues*, which are natural in the early development, representing problems found in developer-system relation and *technical deficiencies* which represent problems of system-specification relation (i.e. bugs/carelessness in implementation). Perhaps the most interesting findings and categories were those of developer-work relation. Some of the implemented functions were revealed to impose *problematic changes* in the work i.e. the system implementation will effect on current practices, existing social relations etc. in a way that users considered them possibly harmful and that designers' may have not understood the effect of some system features yet. In contrast, *unexplored design issues*, are "problems" that could impose positive effects and more value to work of users, if some more design and user research effort were put in the future.

Certainly there can be other categories too (e.g. category of positive findings [9]), as this analysis is based on a single case only. However, what makes these and similar categories together with open tasks feasible also in the tendering situation are that these bring us more profound understanding of what modifications need to be done to the system (and how much these may cost), in order to achieve a successful implementation. As buyers we could compare different systems in tendering against these categories. For example, we could count how many required features are missing, inadequately or

incorrectly designed, misinterpreted by the users or are imposing a problematic change. When each problem in each category is assigned to a feature in the system, we can assess the expected (or formally calculated) costs of implementation (versus benefits).

We do, of course, find the problems of system-work relation also when we test the system against the requirements of the RFP, because the RFP represents work. For example, a feature is missing, if it is requested in the RFP, but it is not implemented into a system. However, if the user fails to complete the task, it may be due to missing functions, missing information, inadequate implementations of the feature, misplaced features at certain point of task flow, or due to the features that are misinterpreted by the users. Thus, single user behaviour and measurement of it (task completion) may divide into several causes in the system (i.e. the problem categories). Similarly, efficiency as measured as time on task or number of errors may be due to many problems in different categories. If the systems are designed with usability workarounds for the verification test purposes only, the buyer may end up in situation where higher usability scores are given to the system that has more profound and costly problems. For example, two systems may be equally efficient in the task of sending email (i.e. having e.g. the same amount of interaction steps), yet the other system can miss a feature of subject field (not required by the task!), include an additional interaction step that is found unnecessary for the task at hand, have inadequate sorting and searching capabilities of email respondents, applies unfamiliar 'compose' and 'primary messages' (see www.gmail.com) instead of more familiar terms 'new' and 'inbox' (see MS Outlook) etc. With traditional measures, these differences would emerge only as a weaker user satisfaction and possibly without more detailed information about the type of the problem.

This is where the categorization of usability problem qualifiers differs from standard measures of usability: Categories do not only give scores but point to specific usability/utility issues in the system that are problematic and require improvements. As categories can have implicit relation to specific phases of human-centred design practices (understand, specify, produce), the buyer is able to evaluate the effort needed to fix the system. Features that are missing and denoting problematic changes at work require more comprehensive understanding of the context of use, while the requirements of misinterpreted and inadequately designed features are only re-specified and technical deficiencies are just re-implemented. The more fundamental the design flaw is and the more complex the future design need is, the more weight the problem category could have in the tendering. An evaluation based on problem categories presumably requires more effort by the buyer. However, we

maintain that systems' usability evaluation during tendering would be more comprehensive as it would contain a broader array of perspectives. Moreover, during the agile development project, categories could improve the design influence by steering and informing developers what features need what kind of modifications and what issues need more user research efforts.

The problem classification scheme is originally a result of applying open tasks in testing. When user-defined test tasks (e.g. open tasks) are applied in the system comparison, the buyer can ensure that, for example the missing system features are the ones needed in real work task accomplishment instead of making the comparison against the requirements in the RFP. This would bring the test closer to a validation situation where the usability requirements in the RFP do not need to be as comprehensive, detailed and rigid, but could be set in higher level (e.g. work roles, jobs, business processes to be supported). In order to test all the systems in a similar way by the specific user, the open task could be co-constructed in co-operation with the user (see [10]), after which it could be locked for this particular user. While problem categories help in answering what is wrong with the system from the viewpoint of users and work, categories cannot substitute usability criteria for time and safety critical systems where e.g. time on task is truly meaningful to measure. Open tasks would serve also better when the emphasis is on the overall system fit with the organizational tasks rather than only on collecting user-system performance data. More experimental research is required, in order to know how to best apply the problem categories and open tasks in the RFPs and IT tendering situations. Before that, we need more analytical research on category definitions as well as measuring and valuing them in the tendering process.

3 Conclusions

- Usability workarounds by the IT vendors may emerge in tendering when rigid and detailed usability requirements are set in request-for-proposals.
- Especially, low-level and detailed formulation of usability test tasks directed the usability design of the case towards sub-optimal solutions i.e. workarounds.
- High-level (usability test) tasks described in the RFP and open-ended usability test tasks for the verification and validation purposes around work could provide flexibility for IT vendors and confidence for IT buyers in introducing and selecting the best technology solution.

- Categories of usability problem qualifiers could be exploited as criteria in the system evaluation during the tendering, in order to bring the evaluation closer to future design and development needs of the system.

Acknowledgements

The corresponding author thanks MATTI doctoral programme of University of Turku for the travel grant.

References

[1] Jokela, T., Laine, J., and Nieminen, M. Usability in RFP's: The Current Practice and Outline for the Future. In Human-Computer Interaction. Applications and Services. Springer Berlin Heidelberg, 2013. pp.101-106.
[2] International Organization for Standardization. ISO 9241-11:1998 Guidance on usability. http://www.iso.org, 1998.
[3] Tarkkanen, K. and Harkke, V. Evaluation for Evaluation: Usability Work during Tendering Process. CHI'15 Extended Abstracts. ACM, 2015.
[4] Tarkkanen, K., Koskinen, J. J. and Harkke, V. Distorted Usability Design in Tendering. ETHICOMP'15, Leicester, UK. 2015 (Forthcoming)
[5] Alter, S. Theory of Workarounds. Communications of the Association for Information Systems 34. ACM, 2014. pp. 1041-1066.
[6] Nielsen, J. Usability Engineering. Academic Press, 1993.
[7] Tarkkanen, K., Reijonen, P., Tétard, F. and Harkke, V. Back to User-Centered Usability Testing. In Holzinger, A. et. al., eds. SouthCHI 2013, LNCS 7946, Springer-Verlag Berlin Heidelberg, 2013 pp. 91–106.
[8] Vilbergsdóttir, S. G., Hvannberg, E. T., and Law, E. L. C. Classification of usability problems (CUP) scheme: augmentation and exploitation. In Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles. ACM, 2006. pp. 281-290.
[9] Tarkkanen, K., Harkke, V. and Reijonen, P. Are We Testing Utility? Analysis of Usability Problem Types. In Proceedings of HCI International 2015, Los Angeles, US. 2015 (Forthcoming).
[10] Tarkkanen, K., Reijonen, P., Harkke, V. and Koski, J. Co-constructed Tasks for Web Usability Testing. In Proceedings of the International Conference on Interfaces and Human Computer Interaction. 22 – 24 July 2013, Prague, Czech Republic, IADIS Press, 2013. pp. 79-86.

User Experience Goals as a Guiding Light in Design and Development - Early Findings

Heli Väätäjä[1], Paula Savioja[2], Virpi Roto[3],
Thomas Olsson[1] and Jari Varsaluoma[1]
[1]Tampere University of Technology, Tampere, Finland
[2]VTT Technical Research Centre of Finland, Espoo, Finland
[3]Aalto University, Helsinki, Finland
heli.vaataja@tut.fi

Abstract. User experience (UX) goals are one means to describe user experience requirements and guide the design and evaluation of interactive systems in different application domains. This position paper discusses the results of a pre-workshop questionnaire for participants of a workshop on UX goals and their utilization. The domains of the case studies that participants described vary from workplace to consumer applications and education. Workshop participants defined a good UX goal as something that (1) helps in focusing the design, (2) is measurable, (3) describes positive emotions, and (4) communicates the desired experience. Furthermore, UX goals were considered useful in keeping the focus on important issues during design and development, and providing inspiration for design.

1 Introduction

To design technology that is capable of enabling, promoting and/or demonstrating specific user experience it is important to set *experiential goals* for the system or solution that is designed. Such *goal setting* approach as part of requirements processes is receiving increasing attention in design and development of interactive systems – not only in relation to everyday consumer technology but also when developing systems for work or education.

Due to the personal and subjective nature of experiences, it is not possible to force people to have specific experiences with products or services. However, designers can aim to facilitate specific experiences among the users, in other words, design *for* an experience [6]. Although the intended experience cannot be guaranteed, it will be more likely if designer utilizes the available knowledge, such as conveyed by Experience Design practice [1]. Commonly agreed experience goals can help the project team by "keeping user experience in focus through the multidisciplinary product development and marketing process" [3]. However, the process of defining, communicating and using

these experience goals as requirements in real life design cases has received little attention in Human-Computer Interaction (HCI) research.

When designing for an experience there is a need to define the *experience goals* that concretize the experiences the users are intended to have when interacting with the developed system. Goal is defined as "the aim or object towards which an endeavour is directed" (in Collins English Dictionary – 30th Anniversary Edition, 10th Edition 2009). By UX goals we mean the experiences that a designer intends the designed system to support for the end-users when they use the system in their activities. These experiences of the end-user are the ones that the system design and development aim to facilitate. The user experience goals can be presented as experiential requirements for the design and development. To concretize the UX goals, they need to be operationalized both for design and evaluation purposes. When UX goals are operationalized, they can be mapped to functional and non-functional requirements, and target user experiences (such as system qualities and impacts [4], as well as feelings and emotions [1]). When UX goals are operationalized, so that they can be measured and therefore evaluated by the users or the verbally expressed experiences of users can be compared to the set UX goals, they are called UX targets.

We held a workshop in fall 2012 at NordiCHI2012 conference [5] to bring together practitioners and academics to share knowledge and lessons learned and to explore:

- *How to identify, define, use and draw inspiration from user experience goals throughout the design and development process?*

The questions driving the workshop included:

- What constitutes a good UX goal?
- How UX goals can be identified and framed?
- How to make use of the UX goals in various design and development phases?

Here we report the results from a pre-workshop questionnaire that addressed the participants' projects described in their position papers. The focus was on defining UX goals, characteristics of UX goals, and their utilization.

2 Method

Before the workshop, the authors of the accepted papers were instructed to complete a questionnaire (in English) dealing with the main themes of the workshop (see Figure 1). We aimed to start the reflection about user experience in relation to the workshop themes and to gain an understanding of

the commonalities across the papers. In this questionnaire, the paper authors were asked to reflect on their personal experiences and conceptions concerning the topic, specifically on usefulness of user experience goals.

The link to the questionnaire was sent to each author of each accepted paper. We received nine responses (response rate 56%) from the possible sixteen respondents. One workshop organizer analyzed the answers to the questions prior to the workshop and discussed the analysis results with another organizer to form a consensus on the results. A summary of the results was presented to the participants before the group work in the workshop. The analysis was revised for this paper to a more detailed level by one of the organizers and authors of this paper.

Based on the case described in your paper, or thinking about some other case, please answer the following questions.

1. Describe the domain of the project briefly
2. Which were the most important UX goals utilised in the project
   a. UX goal 1: _____
   b. UX goal 2: _____
   c. UX goal 3: _____
3. How did you come up with those goals? Based on
   a. a user study
   b. given UX target (e.g. from a customer)
   c. brand
   d. literature
   e. theory
   f. standards
   g. common sense
   h. something else, what?
4. Did you consider any other goals during the project? Which ones?
5. With whom did you communicate about the UX goals?
6. Now looking back, how much did the UX goals affect the design solutions in the project?
   a. 5 very much
   b. 4
   c. 3
   d. 2
   e. 1 not at all
7. Based on you own experience concerning UX goals, please complete the following sentences:
   a. In my opinion a good UX goal is _____.
   b. I have found UX goals useful for _____.
   c. UX goals affect design by _____.

Figure 1. Pre-workshop questionnaire sent to the participants prior to the workshop.

3 Results

Application domains. The papers covered three different application domains or contexts:

1. Workplace: remote operation of cranes, learning tool for forklift drivers, ERP system user interface development.
2. Consumer applications: on-line bingo, designing for dogs, and
3. Education: teaching experience-driven design for university students.

Most important UX goals in reported projects. The UX goals reported by the respondents were relatively similar across the projects. Feeling of safety and feeling of control were mentioned as UX goals in more than one project. The other reported UX goals include: *safety in operation, security, sense of control, feeling of presence, stimulation, competence, self-efficacy, reduced effort, reduced mistakes and errors, freedom from pain and distress, freedom to express natural behavior, and comfort*. In addition, various playful experiences were reported in the paper from education domain that described teaching of experience-driven design for post-graduate students. They include *captivation, submission, fellowship, humour, good mood, amusement, and relaxation*. There is a clear difference between the user experience goals of the paper from the experience-driven design course and the other papers. Although there seems to be a clear divide between pragmatic and hedonic goals here, they both can be equally important for user experience depending on the context or application area.

The sources for identifying UX goals. The most common way to identify the UX goals was a user study (7/9). Literature (3/9) and theory (3/9) were both reported by three respondents. A given UX goal, e.g., given by a customer, was reported by two respondents (2/9). In addition, brand, standards, and common sense were each reported by one respondent. In addition, ethical guidelines and a benchmark study were mentioned as sources for identifying UX goals.

Consideration of other UX goals. When asking whether any other goals were considered during the project (question 4), most respondents reported that UX goals had not been stable throughout the whole project. Three different ways of evolving had taken place based on the responses. First, the goals had been made more precise. Second, some goals had been dropped along the way. Finally, new goals, e.g., business goals, were identified during the project. This indicates that the user experience goals and the requirements may evolve and change to a certain extent as the design process progresses.

Communication of UX goals. Respondents reported that UX goals had been communicated widely in the participating organisations. The stakeholders reported included: customers, users (operators, forklift drivers), domain experts, design or UX team, researchers, colleagues, product owner, and management. This can indicate that UX goals can be one means for powerful and important way of communicating the user experience

requirements to the different stakeholder groups in the design and development phases and when justifying the design choices or even the development in the first place.

Assessing the level of effect of UX goals on the design within the project. Six respondents reported that the utilization of UX goals in the design process had had an effect on the design (three answers missing). Respondents were asked to assess on the scale from 1 (not at all) to 5 (very much) the effect of UX goals on the design. This question received six answers, with four respondents evaluating the effect on level four (4), one as very much (5), and one on level 3, averaging to level 4 (much effect). UX goals therefore seem to be important in the design process and can influence the project outcome.

The characteristics and functions of good UX goals. The responses to the sentence completion concerning a good UX goal covered two distinct aspects: characteristics of UX goals, and the functions of UX goals. Good UX goals were characterized as *measurable* in terms of experience (4/9), *clear and precise* (2/9), *broad* enough to allow space for design ideas (2/9), and *measurable in terms of money* (1/9). One of the respondents describes a good UX goal as "*clear, and enables defining a measurable target*".

The functions, i.e., use of good UX goals were described to be *guiding the design* (2/9), *evoking design ideas* (2/9), and most importantly *communication* (5/9). One respondent describes: "*A good UX goal is giving overall guidance for a certain mindset to follow throughout the whole design process*." Another emphasizes that "*a good UX goal is a way to communicate the desired user-experiences to other people*." All in all, the results emphasize the importance of communicative function of the UX goals as part of UX requirements. Furthermore, the measurability of UX goals was raised as an important characteristic, therefore calling for operationalizing UX goals to UX targets.

The usefulness of UX goals. The second sentence completion focused on the usefulness of UX goals. Four categories for usefulness of UX goals were identified in the responses: focus and framing of UX for design (4/9), to ideate and innovate in design (2/9), to communicate the UX (2/9), to evaluate the design (1/9) as well as to build a business case for developing the UX of a system (1/9). Results indicate that UX goals indeed are useful for various purposes and are important especially in the first phases of the design process and for communicating user's experiences that are aimed for.

Effect of UX goals on design. In the third sentence completion the respondents were asked to state how the utilization of UX goals in the design process affects design. Responses were categorized as follows. The UX goals affect design by providing inspiration (1/9) and vision (2/9), and focus for design (3/9), guiding the design process (3/9), and supporting communication

to educate the organization about UX (1/9). As an example, one of the respondents described: "*UX goals affect design by guiding the product development in its different phases. Their design implications in the context environment should be meticulously defined, e.g., according to the gathered domain and user data, and taken as the guiding stars in the design.*" He/she emphasizes the importance and role of UX goals throughout the development process as well as discusses how other collected information is used for defining the goals. The importance for design process becomes evident from these answers.

4 Conclusion

The results of our questionnaire illustrate that UX goals can be used as an important part of UX requirements. They therefore supplement other means of defining and communicating UX requirements. Specified UX goals can have multiple roles and purposes in the design and development process, such as providing focus and framing of UX, provide a source for inspiration, ideation and innovation, supporting communication of experiential goals to the different stakeholders, supporting evaluation, and building a business case.

As the results come from a very limited set of cases, a more extensive study based on real-life design cases could shed more light on UX goals as part of UX requirements and their presentation. Currently we are working on a more recent data set that was collected during a second workshop in NordiCHI2014 that focused on the first stages of experience design, specifically on identifying and choosing experience goals [2]. Our aim is to understand 1) where designers get insight and inspiration for defining UX goals and 2) what means are used for communicating UX goals among stakeholders. Based on the findings we will present a process model and instructions to support practitioners to identify and define UX goals. The proposed model and instructions will be iterated based on feedback to be collected via interviews and surveys from designers and HCI experts in academia and industry. In addition, a survey on identifying, defining and using UX goals is planned.

Acknowledgements

This research was supported by TEKES (UXUS programme 2011-2015 by FIMECC). The authors gratefully acknowledge this support.

References

[1] Hassenzahl, M. Experience Design – Technology for All the Right Reasons. Morgan & Claypool, 2010.

[2] Kaasinen, E., Karvonen, H., Lu, Y., Varsaluoma, J., Väätäjä, H. The Fuzzy Front End of Experience Design. Workshop Proceedings, in conjuction with NordiCHI2012. VTT, 2015. ISBN 978-951-38-8213-6

[3] Kaasinen, E., Roto, V., Hakulinen, J., Heimonen, T., Jokinen, J., P., P., Karvonen, H., Keskinen, T., Koskinen, H., Lu, Y., Saariluoma, P., Tokkonen, H. and Turunen, M. Defining User Experience Goals to Guide the Design of Industrial Systems. Behaviour & Information Technology journal, Taylor & Francis. 2015. DOI: 10.1080/0144929X.2015.1035335

[4] Väätäjä, H. Framing the User Experience in Mobile Newsmaking with Smartphones. Doctoral Dissertation. Tampere University of Technology. 2014. http://URN.fi/URN:ISBN:978-952-15-3270-2

[5] Väätäjä, H., Olsson, T., Roto, V., Savioja, P., eds. UX Goals 2012 Workshop – How to Utilize User Experience Goals in Design? Workshop Proceedings held in conjunction with NordiCHI2012, Oct. 12th, Copenhagen, Denmark. Tampere University of Technology. 2012. http://URN.fi/URN:ISBN:978-952-15-2955-9

[6] Wright, P., J. McCarthy, and L. Meekison. Making sense of experience. In Funology: From usability to enjoyment, edited by Blythe, M., Overbeeke, C., Monk, A. F., and Wright, P. C., 43-53. Netherlands: Springer, 2005.

Improving User Experience through Task Design and Evaluation Metrics in Research Projects

Xiaojun Yuan
College of Computing and Information
University at Albany, State University of New York
135 Western Avenue, Albany, NY 12222
xyuan@albany.edu

Abstract. In UX requirement practice and research, it has been widely acknowledged that usability and user experience should be addressed in every aspect of the product development process. However, challenges exist on how to consistently and explicitly conduct the usability evaluation and improve user experience. In this paper, I will present how we have tried to improve user experience with interactive information systems by employing a consistent task design and a set of evaluation metrics. An integrated interactive information retrieval system is introduced as an example. Task design and user experience metrics used in a user-centered experiment evaluating this system are described. The task design scheme and user experience metrics have been applied in other user experiments to test the effectiveness of some other information systems and they have been proved effective. Successful experience and lessons learned are discussed. Some further thoughts and suggestions on future work are also presented.

1 Introduction

In a study employing integrative approach to requirements analysis in a user-centric design framework, [5] pointed out the lack of consistency among designers could be an important reason leading to the limited usefulness of the designed system. There exist many reasons that may possibly lead to the ill-defined usability/UX requirements, such as constantly changing instructions, vague user needs, budget concerns, and time constraints. In such cases, both stakeholders and Usability/UX researchers/practitioners take the risk of failing the projects. In 2014, the Standish Group [9] found that more than 50% of projects were reported as challenged and over 30% of projects had to be terminated because of major delays, budget concerns, and so on. There are only about 16% software projects were reported successful in terms of time and budget. It has been an urgent issue for user experience researchers to understand how to maximize the possibilities for a system/software project to be able to succeed in subsequent related projects in the field.

Research has shown that the major causes of project failure are requirements inconsistencies and missing user involvement ([1],[6], [9]). As [1], [5] and [7] stressed, user involvement is a critical factor in the product development process. We should get users involved from the very beginning of the project until the end of it: getting their ideas to better understand their requirements and needs before we design the product, getting their experience of the prototype, and asking them to test the final product and give us their feedback. If the product development is an iterative process, our users should be involved in every single stage of this process.

In the user-centered product development process, task design and analysis is a very important step. It has been found in earlier research that task type has an impact on usability and user experience of information systems [11]. Task analysis [3] provides a technique for designers to characterize their understanding of user behavior in completing a task [8]. Besides task design and analysis, another critical step designers should take seriously is selecting appropriate evaluation metrics in measuring user experience and usability. Rogers, Sharp, and Preece [7] gave designers a detailed description of case studies on topics related to user experience and usability requirements.

The paper explores new methods to address the problem of ill-defined UX requirements in the domain of information retrieval systems and applications. It starts with an introduction of an integrated interactive information retrieval system which was tested in a user-centered experiment. Then a discussion about the user experiment details with particular focus on task design and user experience evaluation metrics will follow. The task design scheme and user experience metrics have been applied in the evaluation of other interactive information systems and they have been proved effective.

2 An Interactive Information Retrieval System

In [11], an interactive information retrieval system was introduced. This system was designed based on earlier findings from [10], and can support multiple information seeking strategies in a single system framework (see Figure 1 through Figure 4). A within-subjects experiment was conducted to compare this system with a baseline generic interactive information retrieval (IIR) system, designed to support searching through query specification.

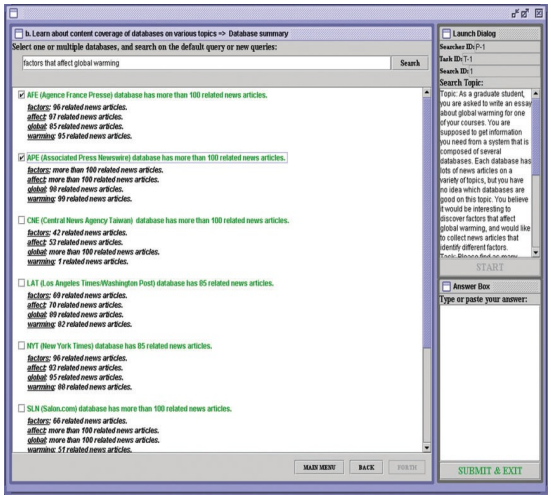


Figure 1. Integrated system, database summary display (c.f. [11])

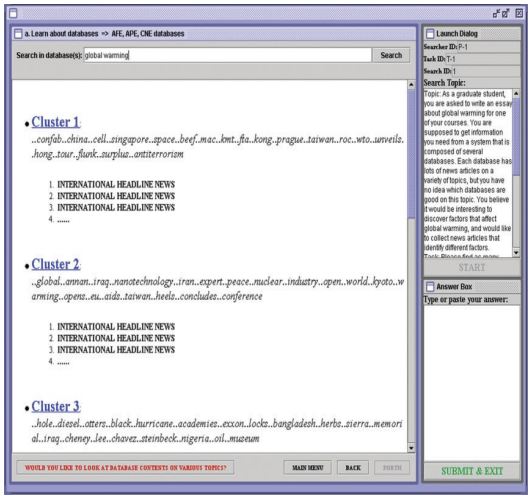


Figure 2. Integrated system, clustered results display (c.f. [11])

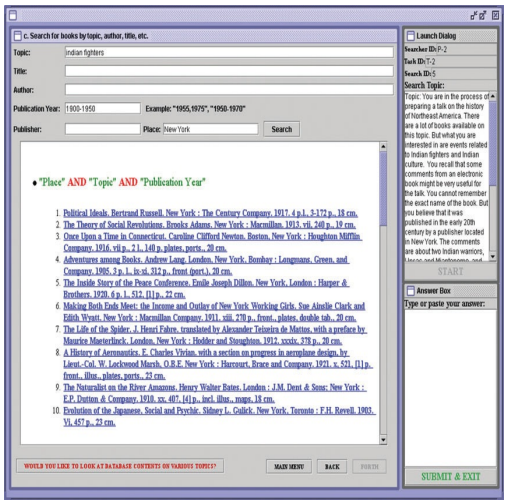


Figure 3. Integrated system, book search results display (c.f. [11])

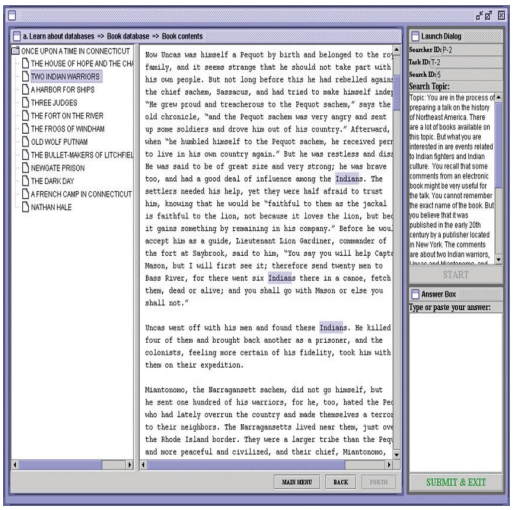


Figure 4. Integrated system, book table of contents display (c.f. [11])

As described in [11], the integrated system begins with a screen containing four options: (a) learning about the databases, (b) learning about the content coverage of databases with respect to a given topic, (c) searching for books on a specific topic, and (d) searching for news articles on a specific topic. Choosing option (a) leads to a screen which lists the names of the databases and the number of documents in each database (see Figure 1). If the user chooses news article databases, the results of the search are displayed as a list of clusters, with highly ranked cluster terms as a label (see Figure 2). If the user chooses a book database, a ranked list of complete citations of the retrieved books will be shown. Clicking on a book leads to a display of the table of contents of the book in a column on the left of the screen. Clicking on one of the items in the table of contents displays that part of the book (see Figure 4). Choosing option (b) leads to a screen which lists 10 queries, each

related to one of the eight test topics, and the two training topics. If the user chooses a news article task, it lists each database and the number of occurrences of each query term in that database (see Figure 1). Choosing option (c), searching for books on a topic, leads to a fielded search/results screen. The user enters values in the fields to perform a query; the results (complete citations) are displayed according to how well they satisfy the Boolean conditions of the query (see Figure 3). Clicking on any citation leads to the display of Figure 4. Choosing option (d) leads the user directly to the search/results screen shown in Figure 2.

3 Task Design

When it comes to real user experiments, sometimes it is challenging to come up with tasks that are similar to real life-time task activities due to the time limitation, lab environment, and diversity of different targeted user groups. Researchers have tried to find a solution to this issue. Simulated Work Task [2] is one of them. Borlund designed the simulated work task situation that can describe users' information need through relevant information, including the source of need, the environment, the information problem, and the goal of the search [2]. We followed the principle of the simulated work task situation model in the user experiment. We categorized the tasks into two different types in order to identify a variety of user behavior patterns. These two types of tasks were designed by following the simulated work task situation model by Borlund [2]. Task type 1 was defined as a task that leads users to engage in scanning, and then searching, while task type 2 was defined as a task that leads users to engage in searching, and then scanning [11].

In the following, we focus on describing how we constructed the tasks. The integrated interactive information system introduced in Section 2 is used as an example.

Before we started evaluating the interactive information retrieval system, we first identified tasks which could lead users to engage in a variety of information seeking strategies (e.g. scanning, and searching) using such a system. We designed the tasks based on our own knowledge of different users' information seeking problems. The specific topics for the tasks associated with book searching were designed so that answers were available to be retrieved from our test collection.

Here is a list of selected task examples we used in the above-mentioned completed user study [11].

3.1 Task type 1: (Finding news articles)

Topic: As a graduate student, you are asked to write an essay about high blood pressure for one of your courses. You are supposed to get information you need from a system that is composed of several databases. Each database has lots of news articles on a variety of topics, but you have no idea which databases are good on this topic. You believe it would be interesting to discover methods that reduce high blood pressure, and would like to collect news articles that identify
different methods.

Task: Please find as many different methods as possible. For each method, please copy the title or link of the article which discusses that method, and paste it to the answer box. For each article that you copy, please type or copy the method(s) that it identifies. If an article discusses more than one method, you only need to copy and paste the article once. If there are several articles which discuss the same methods, you only need to copy and paste one such article.

3.2 Task type 2: (Finding comments)

Topic: You are in the process of preparing a talk on the history of Rome. There are a lot of books available on this topic. But what you are interested in are the wars of Julius Caesar. You recall that some comments from an electronic book might be very useful for the talk. You cannot remember the exact name of the book. But you believe that it was published by a publisher in New York. The comments are about the strategies that Caesar used on the battle field to win the Battle of Pharsalia. You cannot remember the exact comments, but would like to quote them in your talk.

Task: Please find the relevant comments from the book, copy the one best paragraph then paste it into the answer box. Also, please copy the title of the book then paste it to the answer box.

4 User Experience Evaluation Metrics

In the above mentioned user–centered experiment, we measured user experience of an information system as user perception of ease of learning to use the system, ease of use of the system, understanding of the system, and usefulness of the system. These metrics were measured using questionnaires, including post-system and exit questionnaires, all on scales of 1 (low) to 7 (high). The post-system questionnaire elicits user opinions on the system, and

the exit questionnaire gets feedback from users about their experience of the experiment. These metrics helped researchers better understand how users feel about the system, and if they gained satisfactory experience of using the system. Results from the Wilcoxon signed-rank test showed that users found the integrated system to be significantly easier to use, $Z = -2.264$, $p = 0.024$, and significantly more useful, $Z = -2.522$, $p = 0.012$, than the baseline system [11].

It is worthy of mentioning that the above-mentioned metrics were customized in different experiments to better fit the need of system testing. For example, in evaluating the CiteSpace system, a domain knowledge visualization system [4], it was necessary to measure user perception of the various types of knowledge domain visualization graph to better understand if users can understand and use the system well. For this purpose, we added more personalized questions in the experimental questionnaires to get enriched user data for future analysis [12]. This demonstrates that customization or personalization is critical when switching from textual system to knowledge visualization system.

The advantage of using consistent metrics across evaluations of different systems is that researchers can easily compare the findings, and makes it possible to generalize the results. However the disadvantage is that we may fail to collect potentially important data when testing a new type of system. How to design consistent metrics that can be appropriately used in evaluation of various interactive information systems is a critical issue to consider for the UX researchers.

5 Conclusion

This paper described an integrated interactive information system with particular focus on task design and user experience evaluation metrics. Various task types were considered in the task design to better reflect user information needs in using such an interactive information system. The user experience evaluation metrics have been customized for different systems and we are looking to explore how to use the metrics in evaluating mobile applications in the near future. The task design approach and user experience evaluation metrics can be applied to the process of designing and evaluating interactive information retrieval systems. We believe it is critical to know how UX researchers can agree on consistent formats for evaluation metrics, and how such an agreement can better assist researchers in system design and evaluation.

References

[1] Attfield, S., Kazai, G., Lalmas, M., Piwowarski, B. Towards a science of user engagement (Position Paper). WSDM'11, February 9–12, 2011, Hong Kong, China.

[2] Borlund, P. Experimental components for the evaluation of interaction information retrieval systems. Journal of Documentation; 56 (1), 2000, p.71-90.

[3] Card, S., Moran, T. & Newell, A. The Psychology of Human-Computer Interaction. Hillsdale, New Jersey: Erlbaum, 1983.

[4] Chen, S.Y., & Ford, N. Modelling user navigation behaviours in a hypermedia-based learning system: an individual differences approach. Knowledge Organization, 25(3), 1998, 67-78.

[5] Montabert, C., McCrickard, D. S., Winchester, W. W., Pérez-Quiñones, M. A. An integrative approach to requirements analysis: How task models support requirements reuse in a user-centric design framework. Journal of Interacting with Computers, 21, 2009, 304–315.

[6] European software process improvement training Initiative (ESPITI) project: European user survey analysis. Technical Report ESI-1996-TR95104, European Software Institute.

[7] Rogers, Y., Sharp, H., Preece, J. Interaction Design: Beyond Human - Computer Interaction Paperback – June 7, 2011. Wiley; 3rd edition.

[8] Taylor, F., 1991. Scientific Management. Harper & Row, New York.

[9] The Standish Group, 2014. The CHAOS report. <http://www.projectsmart.co.uk/docs/chaos-report.pdf> (retrieved 06.20.15).

[10] Yuan, X.-J., & Belkin, N.J. Investigating information retrieval support techniques for different information-seeking strategies. Journal of the American Society for Information Science and Technology, 61(8), 2010, 1543–1563.

[11] Yuan, X.-J., & Belkin, N.J. Evaluating an integrated system supporting multiple information-seeking strategies. Journal of the American Society for Information Science and Technology, 61 (10), 2010, 1987–2010.

[12] Yuan, X.-J. Investigating the effect of visualization on user performance of information systems. Design, User Experience, and Usability. Web, Mobile, and Product Design: Lecture Notes in Computer Science. Volume 8015, 2013, pp 591-600.

Workshop:
Fostering Smart Energy Applications
(FSEA 2015)

IdleWars: a Pervasive Game to Promote Sustainable Behaviour in the Workplace

Evangelos Tolias[1], Enrico Costanza[1], Alex Rogers[1], Benjamin Bedwell[2]
Nick Banks[3]
[1]University of Southampton
[2]University of Nottingham
[3]Centre for Sustainable Energy
{et2e10, ec, acr}@ecs.soton.ac.uk
Benjamin.Bedwell@nottingham.ac.uk
Nick.Banks@cse.org.uk

Abstract. Energy reduction is one of the main challenges that countries around the world currently face, and there is potential to contribute to this by raising awareness towards sustainability in the workplace. We introduce IdleWars, a pervasive game played using smartphones and computers. In the game, workers' pro-environmental or wasteful behaviour is reflected in their game score, and displayed through eco-feedback visualisations to try and call attention to energy wastage and potentially reduce it. We conducted an initial deployment for two weeks in an open plan office, and observed that the game dynamics works in terms of engaging participants and drawing their attention to the pro-environmental behaviour.

1 Introduction

Sustainability, both in the workplace and at home, has gained increasing attention in the HCI community [2, 10]. Indeed, reducing energy consumption is one of the main challenges that countries around the world are currently facing, and it has been pointed out that raising awareness towards sustainability in the work environment may contribute to this [5]. A large part of this potential is related to energy wastage [4, 13], i.e. equipment being left on when not in use. For example, 9% of workers in the UK reported to never shut the computers down[1]. This energy wastage results not only in additional cost for the industry, but, more importantly, in unnecessary carbon emissions.

We refer to the time that a computer is left on while not in use as "idle time", and we argue that idle time reduction offers a rich opportunity to study interventions to promote energy awareness and behaviour change in the

[1] http://www.1e.com/energycampaign/downloads/PC_EnergyReport2009-US.pdf

workplace. It is worth emphasising that the aim of our work is researching interactive interventions to promote environmental awareness and behaviour change, rather than directly addressing reduction of computer energy consumption. Indeed, the potential energy savings related to computer usage can be considered relatively small (but not miniscule [9] in the context of computer-based work environment) compared to other sources of consumption, e.g. climate control, and there are already commercial solutions available on the market to reduce computer idle time by automatically switching computers off at predefined times (e.g. 1E NightWatchman[2], Cisco EnergyWise Suite[3]). Instead we adopt computer idle time as a *symbol* and as a convenient *research vehicle*, not only because it is easy to recognise as an example of wasteful behaviour, but also because it greatly simplifies the prototyping and evaluation of interactive interventions (as discussed later in the paper). Moreover, idle time is a rather *direct* measure of wasteful behaviour: while increases in energy consumption may be related to increases in productivity [6], idle time is always undesirable.

In this paper we introduce *IdleWars*, a pervasive, competitive game played using smartphones and desktop computers. In our game, workers' pro-environmental or wasteful behaviour is reflected in their game score.

A reduction of wasteful behaviour could be either the result of individual behaviour change, or the result of communal policies, whether at company level or at larger scale. Recently, researchers in sustainable HCI started to question whether eco-feedback visualisations that target individual behaviour change, are an appropriate mean to address sustainability issues [8]. While we largely share such doubts, we argue that when integrated within the broader context of a game for the workplace, feedback visualizations can be instrumental in raising awareness and promoting the proposal and implementation of pro-environmental policies.

2 Related Work

Froehlich et al. provide a review of eco-feedback technology and interventions, discussing the potential benefits of cooperation between the academic fields of HCI and environmental psychology [2]. To date, research related to eco-feedback is mostly focused on the domestic environment [2].

---

2 http://www.1e.com/it-efficiency/software/nightwatchman-enterprise-pc-power-management/

3 http://www.cisco.com/en/US/products/ps10195/

Few papers have addressed eco-feedback in the work environment. Siero et al. [13] investigated behaviour change related to energy conservation in two units of a metallurgical company. Pousman et al. [10] proposed *Imprint*, a system that tracks the documents people print in the work environment, and provides a visualisation of the resources consumed in this way on a semi-public display. Schwartz et al. [12] installed plug-level energy meters in a few offices of a research organisation and observed the reactions of employees through business ethnography. Jentsch et al. [5] presented an energy-saving support system for work environments that leverages a variety of sensors (temperature, electricity, light, contact) to provide workers suggestions about how to act in an environmentally friendly way, however, no real-world evaluation is reported.

Games have also been used with the aim of promoting pro-environmental behaviour, mostly in the domestic context, and targeting children and teenagers [1, 3, 11]. In contrast to this prior work, IdleWars, the game we present here, is a pervasive game designed for adults in the work environment.

The only two games designed to encourage pro-environmental behaviour in the work environment are *Climate Race* [14] and the *Energy Chickens* [9]. Climate Race tracks players' activity in the real world at the room level, through environmental sensors (e.g., switching off lights when not in the office); based on this activity players gain positive or negative points. In the Energy Chickens game, animated characters are used to represent the energy consumed by individual devices in an office (through device level current sensors). Our approach is different from Climate Race and Energy Chickens in three ways: first, it does not require sensors, as activity detection takes place in software on existing office IT infrastructure. Second, in IdleWars the game score is related directly to wasteful behaviour (energy consumed when devices are not in use), rather than the more generic energy consumption. Third, as a pervasive game IdleWars introduces an element of physical action in the real world, beyond screen-based gameplay, with an aim to increase engagement.

The idea of using computer idle time as a proxy for energy wasteful behaviour was originally proposed by Kim et al. [7], who used this measure to investigate two persuasive ambient displays. While their work focussed on individual users, IdleWars uses computer idle time in the context of a game and the online nature of it provides not only personal feedback, but also comparative feedback.

3 Game Design and its Rationale

We started the design process by taking into account the main contrasts between the workplace and domestic environments, to try and best apply lessons from prior work. The first important difference is the lack of incentives: employees generally do not share financial benefits coming from lower energy bills [4]. Another key difference is that workplaces often have a richer social dimension than in a domestic context, not only because generally there are more people in an office than a home, but also because these multiple social groups and layers (e.g. friends, teams, divisions, departments, cross-cutting projects, etc.) may co-exist among workers.

Against this background, we decided to design a game. We believe that through a balance of competition and collaboration games have potential to leverage and influence social dynamics, in a way that can be steered towards pro-environmental behaviour. We decided to focus on wastage around personal *computer usage* for several reasons: first, in the work environment the computer is mostly a personal tool and only its owner has the responsibility of switching it on and off, so it is possible and easy to track *individual behaviour*, in contrast to shared equipment (e.g. from shared printers to coffee machines to corridor lights), for which apportionment would be more difficult or even impossible. Second, in the context of in a midsized office, monitors and computers influence significantly the overall energy consumption [9]. Third, monitoring computers can be achieved *purely in software*, without any additional hardware, therefore keeping deployment costs and installation complexity low and making the system easily scalable.

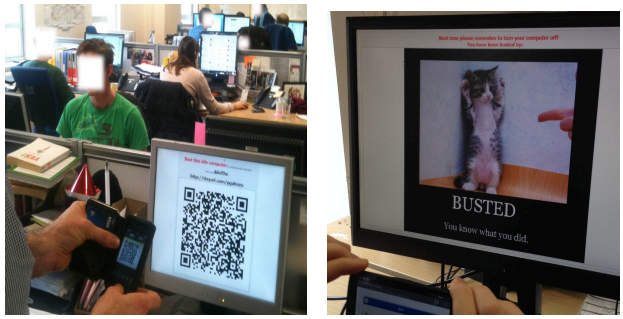


Figure 1. (a) A participant busting the idle computer of another player by scanning the QRCode on the IdleWars screen saver. (b) A busted computer showing the profile picture of the player.

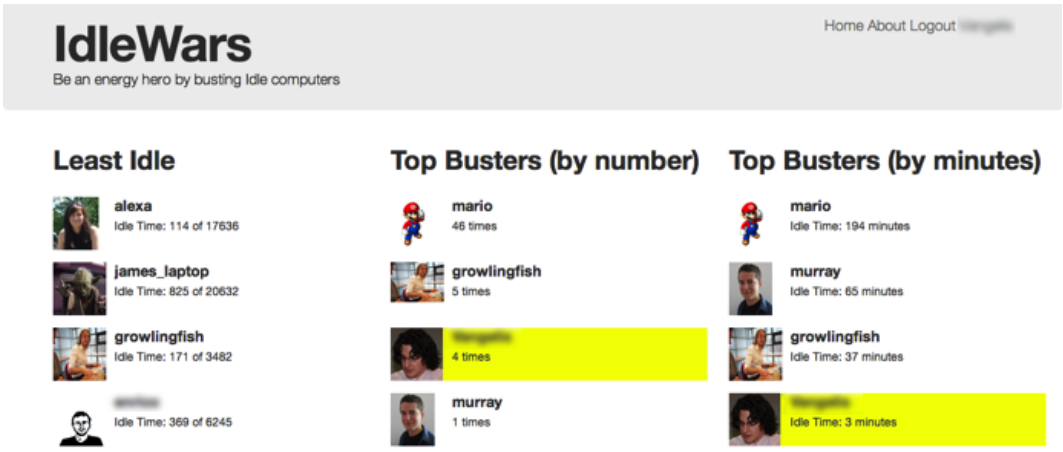


Figure 2. The leaderboard visualisation

IdleWars, the game we designed, tracks the computer status for each player. When no mouse movements or key strokes are detected for more than 5 minutes, the computer is considered inactive, or "idle". In such case, a screensaver appears on the computer screen, showing a QR code, a short url, and an additional alphanumeric code, as illustrated in Figure 1 (a). Any player (other than the computer owner), can then "bust" the idle computer by scanning the QR code with a smartphone, or by manually typing the short url or the alphanumeric code in any web browser (in case a smart phone is not available). Following the busting action, the screensaver of the idle computer changes to show the profile picture of the person who busted the computer, as illustrated in Figure 1 (b). At any point the "owner" of an idle computer, whether busted or not, can close the screensaver and resume the normal operation by moving the mouse or pressing any key. If the idle computer is busted, the owner will see a full-screen profile picture of the player who busted them when they return to their desk. Once a computer has been busted by one player, it cannot be busted by anyone else. The computer based idle time and bust count is presented in aggregate form on the game leaderboard (see Figure 2) via a public display, so our participants were exposed to it throughout the duration of the trial.

4 Initial Evaluation

We report a two-week deployment "in the wild": a medium size organisation, where 20 employees volunteered to participate. Interaction logs were automatically collected by the system, including: idle and active time, bust attempts. Moreover, participants were invited to comment on the trial via an online shared document, and we kept track of emails received from

participants. In total 16 comments were recorded by 11 participants and 9 emails were received from 4 participants. Comments and emails were treated as qualitative data. In total, 14 participants out of 20 busted a computer at least once. We found that the total 37 busting actions took place on just 9 computers, which got busted from 2 to 11 times.

Beside interaction logs, engagement with the game was demonstrated through the choice of profile pictures and emails. During the registration process we described to the participants how the picture would be used. Only six of the participants used an actual photograph as profile pictures, the rest used images of animals, cartoon characters, flowers. One of the participants used an image of a kitten corned with a hand pointing at it, with a message Figure 1 (b). After seeing the game in action, 2 participants emailed us because they decided to change their profile picture to humorous images that included messages for the players they busted (e.g. "I busted you").

We observed that the game sparked discussion within the organisation with regards to computer based energy conservation. The main topic was the trade offs between sleep and hibernation power-saving states in terms of consumption and workflow impact. A few participants commented on technical issues related to the computer power-off, hibernation or sleep, not only in terms of time required to resume work, but also in relation to some specific software applications, which because of interrupted network connection would stop working after resuming.

5 Conclusion and Future Work

In this paper we introduced IdleWars, a pervasive game designed to promote energy conservation in the workplace, in terms of computer idle time. An initial evaluation revealed that the game design works in terms of engaging participants, and raising awareness on energy conservation. It opens the way for further research to investigate the impact of a similar game dynamic other environments such as the home environment. We believe that computer idle time as a measure of pro-environmental behaviour has potential for larger scale, remote deployments, and engagement through online social networks.

Acknowledgements

This work was partially supported by the "Creating the Energy for Change" project (energyforchange.ac.uk) reference No. (EP/K002589/1) and by the Greek State Scholarships Foundation (www.iky.gr), contract No. (2012-ΠΕ-564). We would like to thank the Centre

for Sustainable Energy (www.cse.org.uk) for allowing us to trial our prototype at their premises. The dataset is available at (http://dx.doi.org/10.5258/SOTON/377465)

References

[1] Bang, M., Gustafsson, A., Katzeff, C.: Promoting new patterns in household energy consumption with pervasive learning games. In: Persuasive Technology (2007)
[2] Froehlich, J., Findlater, L., Landay, J.: The design of eco-feedback technology. In: Proc. CHI'10 (2010)
[3] Gustafsson, A., Bång, M., Svahn, M.: Power explorer: A casual game style for encouraging long term behavior change among teenagers. In: Proc. ACE'09 (2009)
[4] Jain, M., Agrawal, A., Ghai, S.K., Truong, K.N., Seetharam, D.P.: "we are not in the loop": resource wastage and conservation attitude of employees in indian workplace. In: Proc. UbiComp'13 (2013)
[5] Jentsch, M., Jahn, M., Pramudianto, F., Simon, J., Al-Akkad, A.: An Energy-Saving Support System for Office Environments. In: Human Behavior Understanding (2011)
[6] Katzeff, C., Broms, L., Jönsson, L., Westholm, U., Räsänen, M.: Exploring sustainable practices in workplace settings through visualizing electricity consumption. TOCHI 20(5) (Nov 2013)
[7] Kim, T., Hong, H., Magerko, B.: Design requirements for ambient display that supports sustainable lifestyle. In: Proc. DIS'10 (2010)
[8] Mankoff, J.: HCI and sustainability: A tale of two motivations. interactions 19(3), 16 – 19 (May 2012)
[9] Orland, B., Ram, N., Lang, D., Houser, K., Kling, N., Coccia, M.: Saving energy in an office environment: A serious game intervention. Energy and Buildings 74 (2014)
[10] Pousman, Z., Rouzati, H., Stasko, J.: Imprint, a community visualization of printer data. In: Proc. CSCW'08 (2008)
[11] Reeves, B., Cummings, J.J., Scarborough, J.K., Yeykelis, L.: Increasing energy efficiency with entertainment media: An experimental and field test of the influence of a social game on performance of energy behaviors. Environment and Behavior (2013)
[12] Schwartz, T., Betz, M., Ramirez, L., Stevens, G.: Sustainable energy practices at work: understanding the role of workers in energy conservation. In: Proc. NordiCHI'10 (2010)
[13] Siero, F.W., Bakker, A.B., Dekker, G.B., Van den Burg, M.T.: Changing Organizational Energy Consumption Behaviour through Comparative Feedback. Journal of Environmental Psychology pp. 235 – 246 (1996)
[14] Simon, J., Jahn, M., Al-Akkad, A.: Saving energy at work: the design of a pervasive game for office spaces. In: Proc. MUM'12 (2012)

Value Sensitive Design Approach to Influence Energy-use Behaviour

Rachel Burrows, Peter Johnson, Hilary Johnson
Department of Computer Science, University of Bath, Bath, BA2 7AY, UK
r.burrows,p.johnson,h.johnson@bath.ac.uk

Abstract. This paper is primarily concerned with relating energy consumption to other information about a person's lifestyle values through sharing and comparing energy-use data with and between others. The goal is to influence choices and in doing so to stimulate sustainable behaviour changes. The study reported in this paper contributes to the design of a web application that aims to achieve this goal. The website allows users to share their lifestyle values and related energy-saving strategies within an online community, and seeks to stimulate sustainable behaviour changes in and through that community.

1 Introduction

A number of interactive technologies have been developed with the aim of reducing energy demand by supporting pro-environmental behaviour e.g. [3,6] They are often based on popular theories of behaviour change, such as value-belief-norm model [14] or action-behaviour-choice [12]. Theories and models can aid designers in understanding what message needs to be presented to the user, and the message framing. Some approaches collect and monitor energy-use information to calculate potential energy-saving opportunities, such as those presented by alternative home heating and cooling programs. This allows people to weigh up the energy costs and benefits of certain usage scenarios and potentially change their behaviour by selecting scenarios with different outcomes and resource implications. Taking a different perspective, some recent work has focussed on models to support user engagement in energy-aware devices [7] this allows user trust to be facilitated through automatic decision making within the system by controlling the degree of autonomy a user can have.

Designing such a system is multifaceted and complex as energy consumption underlies many everyday activities, and also varies greatly between people. Both people's lifestyles and their preferences concerning support to find energy-saving strategies also vary greatly between individuals. As a result, many technologies are reported to result in a lack of long-term engagement and are criticised for interacting with users in ways that are not personalised or relevant to their everyday lives [9,15].

This paper presents participatory design methodology to engage a community of users in the design of this system. Secondly it investigates the issues of personalisation, sharing and privacy of that information. Finally, it investigates the design of interactions and representations to engage and motivate a user community in creating and sharing their strategies, and subsequently using the system to support a sustainable change in favour of pro-environmental behaviour.

2 Our Approach

Our underpinning philosophy is that saving energy is not always the governing or guiding principle around which everyday lives are organised. Everyday lives are organised primarily due to values, around contexts and ways of living. Software systems that aim at influencing behaviour change therefore need to take into account the complex trade-offs that are made to meet the demands and challenges of everyday life while maintaining the values they hold. Of course the value of nature and the environment is a value that many individuals hold [11] and performing activities with the sole purpose of curtailing energy-use may be instrumental in supporting this value. However, it is important to recognise that this is potentially one of many values that an individual or group hold and strive to maintain.

Our work focuses on understanding values as a basis for influencing pro-environmental behaviour. Value Sensitive Design [5] is a methodology for building software applications that takes into account stakeholder values throughout the design process. A value is something that an individual sees as an important aspect of their life. A primary research aim is to understand how people's values can influence long-lasting behaviour changes. From this understanding it is possible to investigate attributes of software systems to influence people to reduce their high carbon, energy consumption behaviours.

People have many values. They are used to select and justify activities, and to evaluate artefacts and events (including other people and themselves)[11] The values people hold are considered as important influencers and drivers for their specific energy related behaviours [10] Those values may be articulated in rather generalised forms by collective terms. Existing research into the structure and content of values can be seen in work by Schwartz [11] who presents a set of 10 universal value types including security, conformity and tradition among others. Similarly, work by Rokeach [10] also presents work on collective values including a comfortable life, social recognition and wisdom. Each of these types contains subtypes of instrumental motivational

values; for instance the value type of security contains a set of 7 instrumental subtypes such as family security, healthy, social order, and so on.

However, these generalised value forms take on real meaning and influence at a much more personal and individual level. Moreover, the values for an individual are constructed and operationalised through the connections they have with the terms and contexts of their everyday life.

3 Study Design

The objective of this focus group session was to investigate the design of web-based interactions and representations to engage the user community. The activities to be supported were those of searching for, creating and sharing their strategies and subsequently using the system to support a sustainable change in energy behaviour. Thus there are two fundamental activities that are supported by the website that may facilitate users to reduce their energy consumption. These are 1) allowing individual users to create their own energy-saving strategies for the user community and 2), allowing users to search the community energy-saving strategies. Through the activities of this focus group participants generated prototype designs for the web application. It is important to note that the types of users that would be members of the online community remained an open design question. In other words, the participants were free to suggest who would be members of their online community within their prototype designs.

3.1 Sample

The sample of 8 participants, some of whom had been involved in previous sessions, were chosen to represent a variety of energy users in the UK. They were paid £35 in vouchers to participate. The participants had different levels of familiarity with web technology, lifestyles, age ranges, and employment status. Participants carried out two design sessions in pairs, one as the interviewer and one as interviewee. Partners and roles were swapped between sessions. The first design session focussed on contributing their own strategies on the website. The second design session focussed on searching and finding energy-saving strategies of interest on the website.

3.2 Procedure

The methodology follows the a series of participatory design [8] techniques proposed by the D.School at the Institute of Design at Stanford University [4]

The method is grounded in first gaining an understanding of user values and lifestyles, and progresses through to prototype development. System features created during the prototyping phase are traceable back to the design problems they were addressing and also the specific user characteristics that motivated them in the first place.

3.3 Design Workbooks

For each design session, each pair was provided with a workbook containing 9 design questions. They were also given a collection of rapid prototyping materials which included pens, generic user interface shapes, stickers, scissors, colours, highlighters, UI outlines of tables, mobile phones, laptops. The workbooks contained 9 questions, each within one of the following design phases.

Understand (Q1 and Q2): Interviewer was to gain a deeper understanding of their lifestyle, values and also user-specific characteristics that may motivate them to share their strategies and adopt the strategies of others.
Problem Statement (Q3 and Q4): Each pair specified problems that are potential barriers to engagement with the system.
Ideation (Q5): Each pair sketched a selection of ideas that would solve their previously defined problem.
Iterate (Q6 and Q7): Each pair refined their ideas into a single prototype.
Build and Test (Q8 and Q9): Each pair consolidated their design idea into a final paper prototype. They evaluated this prototype.

4 Results

The participants worked in pairs and produced a total of eight design books - four addressing designs to support strategy searching and four to support strategy creation. All eight designs were grounded in an understanding of the participants' lifestyle values and energy behaviours.

The lifestyle-value-driven approach informed many aspects of the UI design. The design ideas were conveyed in drawings of UI designs and annotations on those designs. This included requirements that related to; 1) issues of security and privacy, 2) how content on the website is organised and framed, and 3) how they would want to interact with the website. Final prototype solutions generated by the pairs varied. For instance, one pair drew a process diagram representing the flow of interaction with the website. Others used the given screen templates together with the prototyping materials to

illustrate layout details, while others used the materials to convey certain aspects of functionality. Drawing from the resulting focus group workbooks, lists of requirements, user characteristics and general comments about system preferences were constructed. The following outlines some findings from the workbook answers.

Results from the workbook answers showed how multiple participants had ideas about what kinds of energy-saving strategies they would want to search for and find. Four of the eight workbooks found answers that indicated that it was important to find energy-saving strategies that were relevant to their lifestyle and values. For instance, one pair wrote that they would be interested in sharing strategies with other users where they had ``something in common''. Other participants wrote how they would be interested in finding energy-saving strategies that had ``personal relevance'' or that the website content could be ``refinable by circumstance''.

All the workbook contents suggested a variety of different characteristics of strategies that would motivate them to read a particular strategy. Some of these characteristics were expressed through the drawing of UI menu headings to show visually how the content should be divided into separate parts of the website. Additionally, some characteristics were expressed through written answers. There was no consensus to these characteristics. Some workbook examples of strategy characteristics that would be of interest were: Monetory Benefit; Immediate benefit; Little effort; Funny; New and Novel; Tried and Tested; Recommended by Friends (highly voted for). Due to the lack of consensus on strategy characteristics that users would be interested to read about, accessing content through fixed menu items may not be as effective as allowing strategies to be organised by user-driven labels.

The main motivation for users to share their own strategies was the impact it may have upon their online community. For example, one pair wrote how combined efforts with others would motivate an effort to share energy-saving strategies; evidence of ``reciprocity'' within the online community was the key to maintaining engagement. Another pair wrote that a benefit of an online community for communicating strategies is that it has the potential to reach many people; they wrote that ``spreading the message'' would be a motivator.

Some participants expressed concern about what information they would be sharing, which may act as a barrier to engagement. This concern was expressed in many ways. For instance, some participants stated how a barrier to engagement may be due to them being self-conscious about what they were revealing, while other participants explicitly stated that that they would be concerned about security.

5 System Design Implications

The design book results from section 4 influence many aspects of the website design. The main challenge was to converge multiple design ideas into one solution that addressed the design issues raised by the participants. This is not straightforward due to the requirement conflicts that became apparent between participant designs and the website goals, and also between individual participant designs themselves.

The website should allow users to visualise how relevant energy-saving strategies are to their values. The visualisation will simultaneously address both participants prototype suggestions of the website content to be personalised at the same time as communicating strategies in a way that will influence users to adopt the strategies they find. For instance, users may be more likely to adopt strategies that support the continuance of highly valued aspects of their lifestyle. Secondly, viewing strategies from other users with similar lifestyle values may lead to a social influence effect causing users to adopt strategies of other users within the online community.

The website should additionally allow users to add further information, in the form of user-generated labels to energy-saving strategies within their online community. This would allow for the website content to be searchable according to those user-generated labels. This design decision was driven by the large amount of variability in the types of strategies that users would be motivated to read about. These categories are typically not currently used in current energy-saving applications, which often use dwelling-related categories of energy-use. For example, existing systems may present users with energy information according to the type of energy that is being used (gas, electricity etc) or alternatively the room in the home (kitchen, bathroom). It does, however, allow users to self-organise website content according to what is of interest which can in turn be used to evaluate alternative behaviour change campaigns and policies.

The website should provide users with the option to vote and comment on published energy-saving strategies to allow for users to prioritise visibility of popular strategies. This design decision is in response to the fact that visible feedback from strategy readers is important to show evidence that the contribution from an individual user is of benefit to the larger community. It is also beneficial to rank the strategies and give priority to high-quality content. This interaction was also justified by two participants who reported to be wary about interacting with the website. These interactions may serve to build confidence in users by providing evidence that their contribution is useful to others.

6 Conclusion and Future Work

These design decisions are an important step towards the first implementation of the website. User values have informed website design in two ways. Firstly, lifestyle values and associated energy-saving strategies are integral to the structure and content of the website representations and interactions; they are a mechanism for searching and creating the user-generated content. Secondly, user values have also informed the way in which searching and creating strategies is achieved. For instance, values of security, community, privacy and reciprocity are also integral in the design of website in order to maintain user engagement.

The next steps are to get that community of users using the website and evaluate both the design for its HCI and the approach for its ability to enable strategy generation, strategy sharing and strategy take up by the community of users. This would then allow us to further evaluate if it produces energy-related behaviour changes derived from the common and different lifestyle values of individuals and the creation and adoption of community-generated energy strategies.

Acknowledgements

Eviz Project. The work reported in this paper is funded by the Engineering and Physical Sciences Research Council (EPSRC) under the Transforming Energy Demand in Buildings through Digital Innovation (TEDDI) (grant reference EP/K002465/1).

References

[1] Burrows, R., Johnson, P., and Johnson, H. Influencing behaviour by modelling user values : Energy consumption. In PERSUASIVE'14 2nd Intl. Workshop on Behaviour Change Support Systems (2014).

[2] Carter, S., and Mankoff, J. When participants do the capturing: the role of media in diary studies. In Proc. of the SIGCHI Conf. on Human Factors in Computing Systems, CHI '05, ACM (NY, USA, 2005), 899–908.

[3] Costanza, E., Ramchurn, S. D., and Jennings, N. R. Understanding domestic energy consumption through interactive visualisation: a field study. In Proc. of, UbiComp'12, ACM (NY, USA, 2012), 216–225.

[4] DSchool. Design thinking website http://dschool.stanford.edu/dgift/. accessed 10th Feb 2015.

[5] Friedman, B., Kahn, PeterH., J., etal. Value sensitive design and information systems. In Early engagement and new technologies: Opening up the laboratory, N. Doorn, D. Schuurbiers, I. van de Poel, and M. E. Gorman, Eds., vol. 16 of Philosophy of Engineering and Technology. Springer Netherlands, 2013, 55–95.

[6] Ham, J., and Midden, C. A persuasive robotic agent to save energy: The influence of social feedback, feedback valence and task similarity on energy conservation behavior. In Social Robotics, vol. 6414 of LNCS. Springer Berlin Heidelberg, 2010, 335–344.

[7] Hammer, S., Winer, M., and Andr, E. Trust-based decision-making for energy-aware device management. In User Modeling, Adaptation, and Personalization, vol. 8538 of LNCS. Springer International Publishing, 2014, 326–337.

[8] Muller, M. J., and Kuhn, S. Participatory design. Commun. ACM 36, 6 (June 1993), 24–28.

[9] Pierce, J., and Paulos, E. Beyond energy monitors: interaction, energy, and emerging energy systems. In Proc. of the SIGCHI Conf. on Human Factors in Computing Systems, CHI '12, ACM (New York, NY, USA, 2012), 665–674.

[10] Rokeach, M.: The nature of human values. Free press (1973)

[11] Schwartz, S.H.: Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. Advances in Experimental Social Psychology, vol. 25, pp. 1 65. Academic Press (1992)

[12] Shove, E. Beyond the abc: climate change policy and theories of social change. Environment and planning. A 42, 6 (2010), 1273.

[13] Simon, H. A. A behavioral model of rational choice. The Quarterly Journal of Economics 69, 1 (1955), 99–118.

[14] Stern, P. C., Dietz, T., Abel, T., Guagnano, G. A., and Kalof, L. A value-belief-norm theory of support for social movements: The case of environmentalism. Human Ecology Review 6, 2 (1999), 81.

[15] Strengers, Y. A. Designing eco-feedback systems for everyday life. In Proc. of the SIGCHI Conf. on Human Factors in Computing Systems.

Visualizing and Gamifying Water and Energy Consumption for Behavior Change

Isabel Micheel[1], Jasminko Novak[12], Piero Fraternali[3], Giorgia Baroffio[3], Andrea Castelletti[3], Andrea-Emilio Rizzoli[4]
[1]European Inst. for Participatory Media, [2]Univ. of Applied Sciences Stralsund, [3]Politecnico di Milano DEIB, [4]IDSIA USI / SUPSI
[1]Berlin, [2]Stralsund, Germany, [3]Milano, Italy, [4]Manno, Switzerland
{i.micheel, j.novak}@eipcm.org

Abstract. This paper considers the structural similarities in approaches and lessons learned in the development of applications for behavior change in water and energy saving. We show how the domains of water and energy are related and propose a first set of design guidelines for building such solutions, especially regarding visualization and gamification of water and/or energy consumption. We exemplify how such guidelines can be applied with the designs and prototypes of a gamified application for water saving behavior change from our SmartH2O project. Based on feedback from user and stakeholder workshops and online discussions, we discuss how the initial design guidelines synthesized from the literature have been refined. In a next step, we will validate them by deploying the implemented prototype in real-world trials with several thousand smart-metered households in the UK, Switzerland and Spain.

1 Motivation

Raising consumer awareness and stimulating behavior change in the use of natural resources in different domains (e.g water, energy) has become an important research challenge. Supporting behavior change can be fostered by raising consumers' awareness [2] at individual and collective level, by providing actionable recommendations fitting the consumer's context and relating incentive models to consumption habits and the consumer's community of reference. Existing approaches to water and energy consumption differ in some aspects but also exhibit many common traits and findings: from common challenges and incentive schemes to impact potential (see Section 2). Though this makes sense intuitively, since both resources are often consumed together (e.g. hot water), only few attempts have compared the approaches from both domains. We have performed such an analysis with a specific focus on visualization and gamification, and extracted common elements as general guidelines for designing applications that aim at raising awareness and stimulating behavior change in resource consumption. By

applying them in designing applications for water saving in the SmartH2O project [17], we refined them with feedback from users and stakeholders.

2 Lessons from existing work

A number of water and energy conservation applications for consumers employ visualized consumption feedback and gamified social interactions to motivate people to adopt more sustainable lifestyles, with various level of success [6][15]. Common approaches can be identified, based on two main shared challenges:

- How to present consumption information and convey its meaning to users (increase awareness)?
- How to enable and motivate consumers to change their consumption (induce & sustain behavior change)?

The first challenge is often addressed by visualizing consumption information. Common approaches have been data-oriented (e.g. bar or pie charts [6][12][17]), closely connected to the real consumption context (e.g. floor plans [6][12]), metaphorical (e.g. traffic lights and gauges [12][15]), or playful and ambient [10][11], often connected to nature or animal habitats (eco-visualization) [6][8][15]. To visualize consumption effectively, it can be broken down, e.g. temporally or by events and type of consumption. A study on visualizing water consumption identified four eco-feedback design dimensions that should inform visualizations of water consumption: data and time granularity, comparison, and measurement unit [6]. It has shown that study participants valued data granularity at individual fixture level or fixture category [6]. The need to visualize consumption per appliance is also highlighted in energy research [8][9], suggesting that it facilitates long-term sustainable behavior [8].

The use of benchmarks for comparison is important, as they allow users to judge whether their consumption is “normal”, excessive, or economical [15]. Comparing consumption is essential for helping consumers to understand it. A user study on water consumption found out that the provision of different ways of comparison (self-, goal- and social comparison) was highly appreciated by participants, especially self-comparison of a household’s current vs. historical consumption [6]. Goal-comparison was most valued for self-set consumption goals, and least for goals set “top-down” by suppliers or local governments. Social comparison was also popular, especially with similar households and geographic neighbors [6]. In the energy domain, similar kinds of comparisons have been considered if under different terms, e.g. historical for self-comparison, normative for comparison with other

households and 'social' comparison for comparison against others "in their collective social setting", e.g. individuals in the same household [1]. Most studies argue that consumption comparison is an effective means for stimulating behavior change, with some controversy over whether social comparison [5] or historical comparison [9] has a greater effect, but a user-based validation of different kinds of comparisons is largely lacking.

To address the second challenge of stimulating consumers to change their behavior, the provision of action-oriented tips for consumption reduction is a common strategy in both domains [3][6][7][9][14]. Actionable tips are needed, as the visualization of consumption information alone doesn't in itself provide practical hints on how to improve it [15]. Consumption behavior tips can be more general or contextualized (e.g. concerning overall consumption or specific consumption areas) [7][9]. Most of them are not personalized, i.e. not adapted to the characteristics of a household or a consumer and their consumption behavior.

To motivate consumers to act upon presented consumption information and tips, gamification is increasingly applied. Studies on gamifying energy consumption have shown e.g. that real prize-like rewards can be effective incentives [10][14] but also that gamified social interaction can foster better behavior, through both competitive and cooperative approaches [3][6]. Whether competition or cooperation works better hasn't really been validated yet, but some work recommends making competition optional and stressing collaboration instead [6]. Similarly, while playful designs can be engaging, special care needs to be taken to adapt visual style to semantic meaning (e.g. more visually appealing the more is saved) and to focus on portraying actionability [6][11].

3 Preliminary design guidelines

Based on this analysis, we extracted a set of preliminary design guidelines for resource consumption awareness applications (see Table 1). They summarize main aspects for the effective design of applications applying visualization and gamification to support reduction of natural resource consumption (water, energy) by raising awareness and stimulating behavior change.

Table 1. Design guidelines synthesized from related work

Design guideline (DG)	Objective
a) Visualize consumption in an understandable form	Raise individual awareness by conveying meaning
b) Visualize specific dimensions of consumption	Raise individual awareness by conveying meaning
c) Visualize comparisons to relevant references	Raise individual & collective awareness by (social) comparison

d) Provide consumption tips	Enable individual behavior change
e) Gamify consumption	Motivate individual behavior change
f) Stimulate social collaboration & competition	Motivate individual and collective behavior change

To explore their applicability and further refine them, we have applied these guidelines in developing a web and mobile application combining smart water meter consumption readings with visualization and gamification within the SmartH2O project [17]. In the next section, we present first insights from our design cycle, in which we have developed a series of mockups and visual prototypes and collected feedback from stakeholders and target users in workshops and an online discussion space.

4 Prototyping & user feedback

We are developing our application according to these guidelines in an iterative user-centered design approach that combines user-driven needs of different stakeholders (user pull), including water consumers and utilities, and state-of the-art technological advances (technology push) [13]. As a first outcome, visual prototypes accompanied by narrative user stories to contextualize the concept were developed and adapted to user feedback in three main iterations.

Raising awareness through gamifying consumption

The first visual prototype (see Figure 1) depicts both a web and mobile version of a gamified application, which is connected to a user's smart meter to measure his household water consumption. It contains a basic visualization showing aggregated consumption, which enables self-comparison, e.g. by providing metrics such as averages and peak consumption at different time-granularities, goal-comparison, and comparison with the average consumption of one's neighborhood. The application is gamified, such that all user actions, including providing information about their household or reading and implementing water saving tips, as well as water saving efforts, are translated into virtual points.

Figure 1. First visual prototype of the gamified portal

With these, users can earn reputational badges for different types of actions, and they can redeem real rewards provided by the utility or external sponsors. To stress the actionability of the approach, users are suggested concrete actions to increase their point score and reach the next badge. Based on their activity and total points, they are also ranked with others on a leaderboard to stimulate social comparison and competition with other households (e.g. from the same town or with similar characteristics). A neighborhood map shows households that are geographically close and that are also participating in the water saving efforts. To facilitate collaboration, users can form and join water saving teams (Figure 1, right). Team members benefit from each other's actions and can work towards common water saving goals. The prototype was discussed in a workshop session with 30 local residents in a Swiss municipality. A main concern from participants was that they wanted to know exactly what benefit they would get out of such an application, beyond virtual points and saving water to help the environment. An additional feature appreciated by most was the idea of warning alerts in case of e.g. leaks, overconsumption or upcoming shortages and water quality issues. Thus, an important finding was that, to reach a larger audience, more pragmatic users should be considered in addition to those that would embrace hedonic, playful elements like badges, competition and collaboration. But, workshop participation itself was successfully gamified through a raffle which gave away water saving gadgets. This was an effective motivation even for those hesitant towards gamification, which indicates that *gradually introducing new, pragmatic users to the idea of gamification* with real rewards could engage them eventually.

Differentiating pragmatic and hedonic scenarios and users

As a response to this deeper understanding of the target users, the different envisioned features were distinguished more clearly to allow *separate views for pragmatic and hedonic users*. Two versions of the portal prototype are implemented: a basic version that targets more pragmatic, data-affine users by focusing on consumption visualization and practical water consumption alerts and tips and an advanced version which introduces gamification and social features in addition to the central visualization to target more hedonic users (see Figure 2). In addition, to increase pragmatic value, personalized feedback and water saving tips will be provided based on analyzing consumption behavior, to identify consumer classes with shared consumption patterns, household and personal characteristics [16].

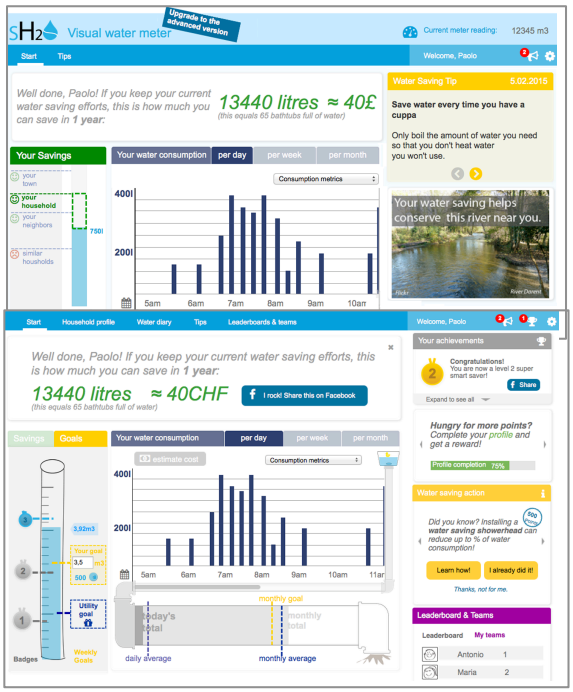


Figure 2. Visual prototypes: basic portal (l), portal with gamified visualization (r)

To differentiate between users that can be motivated by gamification, and those more likely to use social features (optional to address privacy concerns of users who may not want to share information with others), player types have been considered [3]. Table 2 distinguishes target user and player types.

Table 2. Target user attributes for the application versions

Application version	Tech. affinity	Data affinity	Playfulness / Player type
Basic portal: visual water meter	Low	High	Low
Advanced portal: gamified meter	High	Low	Achiever, explorer
Advanced portal: social meter	High	Low	Competitor, socializer

Feedback from a second workshop in the UK with 11 participants confirmed the finding that separating the three functionalities is likely to improve user acceptance. While the workshop in Switzerland was attended by residents of all ages, participants in the UK were generally younger (average age ca. 35), more familiar with the concept of gamification, and stated that environmental concern would be one of their strongest motivators (*value-based intrinsic motivation*). Most considered the idea of a mobile app to monitor their water consumption very appealing, especially for quick access and alerts. The notion of competing against neighbors was not found particularly stimulating. However, they would consider competition with family and friends whom they would trust more and would feel more

connected to (*embedding in relevant social context*). On the other hand, the idea of pursuing joint goals as a community and of receiving collective rewards sparked enthusiastic responses, e.g. "Could be a good experience, you get to know your neighbors better, especially in urban areas people don't know their neighbors any more, nice to be a kind of community." (*in-group collaboration, intra-group competition*).

Designing actionable consumption visualizations

In a next iteration, a more detailed visualization model was developed, addressing the design guidelines and different dimensions discussed in Section 2 in more detail, e.g. data and time granularity, different types of consumption comparison and different measurement units. Consumption information is visualized at different levels of detail in a way that maps abstract metering data into a form understandable for users, raises consumer awareness and enables them to act upon it accordingly. Overview visualizations provide users with simple messages regarding their water consumption. One widget, e.g., visualizes consumption savings (or water wasted) compared to others (similar households, neighborhood, households in your town), addressing social comparison aspects of the DGs; an ambient eco-visualization targets environmentally conscious users, showing nature reserves affected by users' water consumption (Figure).

To visualize consumption in more detail, a water pipe metaphor is used, which conceptually connects to the real consumption context but shows information as a simple bar chart (see Figure , left). It displays the total consumption for different time intervals, compared to the average consumption (self-comparison). Monthly goals, set by users themselves or their water utility, are also indicated (goal-comparison). By breaking down smart meter data further with disaggregation algorithms [16], consumption percentages for fixture types and end-use events are also visualized (see Figure , right). End-use events are detected automatically or edited manually, e.g. for corrections. Events are visualized as fixture icons corresponding in size to the amount of water consumed. *Actionable consumption behavior tips are embedded in the visualization* when overconsumption is likely or has occurred, to enable users to change their behavior accordingly.

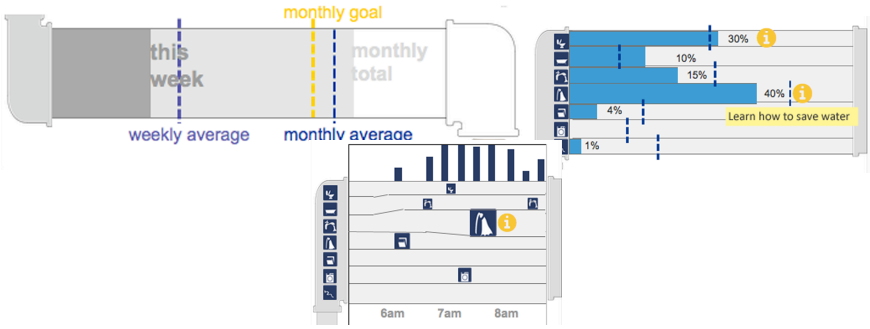


Figure 3. Pipe metaphor: aggregated (left), by fixture (center), by end-use events (right)

Household consumption is visualized in more detail with a bar graph at daily, weekly or monthly level, with metrics like consumption peaks. Embedded in the third prototype (see Figure 3), the visualization was discussed with different stakeholders, including water consumers, utility staff and researchers, in a moderated online visual discussion space for two weeks. Overall, discussion participants showed high interest in the SmartH2O concept and prototypes (the discussion attracted 80 new members and 50 comments during the period). A key user comment on the visualization was e.g. "I would suggest something more 'basic' with an option to switch to more detail/ or advanced mode for the geek ones" *(present interactive layered visualization; simplest by default)*. And, while we considered analogies for the amount of water saved, feedback from the discussion suggests that showing consumed water may have a bigger impact as the numbers are larger: "You could try using analogies like how many olympic-sized pools can you fill with the water consumed. I believe that you can provide a 'productive' shock to the consumers this way" (*visualize different measurement units and metaphors for consumption and saving*).

5 Conclusion and future work

Based on the preliminary experiences from the described design cycle and user workshops, we have refined the design guidelines for visualizing and gamifying resource consumption (water, el. energy) extracted from literature (see Table 3).

Table 3. Refined design guidelines from SmartH2O experience

DG	Refined aspects from SmartH2O design cycle & user workshops		
a	Present interactive layered visualization (simplest by default)	**c**	Goals should be related to concrete actions users can perform
	Use visual metaphors relating to user's consumption context	**d**	Feedback on consumption should be action-oriented and include saving tips embedded in the visualization
b	Present separate views for less vs. highly data-affine	**e**	Real rewards should engage even more pragmatic users

	users		
	Overview of consumption should trigger awareness, and detailed information should point out concrete actions		Separate views for pragmatic & hedonic users should be considered
	Units & analogies should illustrate consumption & savings	**f**	Common goals have the potential to bring e.g. neighbors closer
	Showing consumption rather than smaller savings can raise awareness		Both are promising for different users embedded in the relevant social context: in-group collaboration, intra-group competition

While these design guidelines contain important aspects, they so far reflect our exploratory enrichment of literature findings in which only some of the studies included longitudinal field evaluations (and none large-scale usage). Thus, we are implementing the two application versions as presented, and validating them with several thousand households in trial areas in the UK, Spain and Switzerland over the next two years. As one of the outcomes of this validation, we hope to extend the guidelines and formulate design patterns.

References

[1] André, A., Buhling, R., et al. Evaluating the effectiveness of visualizations for comparing energy usage data. In FSEA 2014, pp. 9–14.

[2] Darby, D. The effectiveness of feedback on energy consumption. Report. Env. Change Inst., Univ. of Oxford, 2006.

[3] De Luca, V. and Castri, R. The social power game: A smart application for sharing energy-saving behaviours in the city. In FSEA 2014, pp. 27–30.

[4] Ferro, L. S., et al. Towards personalised, gamified systems: an investigation into game design, personality and player typologies. In Proc. IE2013. ACM.

[5] Foster, D., Lawson, S., et al. Wattsup? Motivating reductions in domestic energy consumption using social networks. NordiCHI 2010, ACM, 2010.

[6] Froehlich J., Findlater L., et al. The design and evaluation of prototype eco-feedback displays for fixture-level water usage data. In CHI 2012, pp. 2367.

[7] Gamberini, L., Corradi, N., et al. Saving is fun: designing a persuasive game for power conservation. In Proc. ACE 2011, ACM, 2011, pp. 16.

[8] Gustafsson, A. Power explorer: a casual game style for encouraging long-term behavior change among teenagers. Proc. ACE 2009, pp. 182–189.

[9] Jacucci, G., Spagnolli, A., et al. Designing effective feedback of electricity consumption for mobile user interfaces. PsychNology J. 7(3): 265–89, 2009.

[10] Kappel, K. and Grechenig, T. "show-me": water consumption at a glance to promote water conservation in the shower. In PERSUASIVE 2009, pp. 26:1.

[11] Laschke, M., Hassenzahl, M., et al. With a little help from a friend: a shower calendar to save water. In CHI'11 Ext. Abstracts. ACM, 2011.

[12] Monigatti, P., et al. Power and energy vis. for the micro-management of household electricity consumption. In Proc. AVI 2010, pp. 325–28.

[13] Novak, J., et al. histoGraph – A Vis. Tool for Collaborative Analysis of Historical Social Networks from MM Collections. In Proc. iV2014, Paris.

[14] Peham, M., Breitfuss, G., Michalczuk, R. The 'ecoGator' app: gamification for enhanced energy efficiency in Europe. In Proc. TEEM '14. ACM, 2014.

[15] Rist, T. Towards a more responsible use of energy through visualization of energy data. In FSEA 2014, pp. 9–14.

[16] Rizzoli, A.E., et al., The SmartH2O project and the role of social computing in promoting efficient residential water use: a first analysis. In iEMSs 2014.

[17] SmartH2O project. EU FP7, No. 619172, 2014-17. URL: smarth2o-fp7.eu/

[18] Valkanova, N., et al. Reveal-it! The impact of a social visualization projection on public awareness and discourse. In CHI 2013. ACM, pp. 3461.

Promoting Energy-Efficient Behavior by Recommendations based on Energy Cultures

Stephan Hammer*, Fabian Segmüller*, Birgit Lugrin+, Elisabeth André*
*Human Centered Multimedia, Augsburg University,
Universitaetsstr. 6a, 86159 Augsburg, Germany,
{hammer,andre}@hcm-lab.de
segmueller@student.uni-augsburg.de
+Department of Computer Science, Würzburg University,
Am Hubland, 97074 Würzburg, Germany,
birgit.lugrin@uni-wuerzburg.de

Abstract. Enhancing people's energy behavior is an important research topic. We aim to support people by recommending specific energy-saving actions when they have the opportunity to save energy. To persuade people to execute the actions we incorporate findings of models related to behavior change into the system's decision making process. In this paper, we consider how the idea of Energy Cultures can be integrated into a recommender system and present results of an online survey that investigates the effect of users' Energy Culture on their preferences for energy-saving actions.

1 Introduction

For several decades research investigates how people's energy demand can be reduced [2]. Developed approaches include the provision of detailed feedback on people's energy usage [5], persuasive games [6], and systems, such as smart thermostat controls [8], which automatically measure and control the energy consumption. However, people still lack accurate and easily accessible information on how they could achieve potential savings through own actions [7]. Benders et al. [1] showed that the best results can be achieved if people receive useful and personalized advices that they are able to perform, and that they can trust. Ford et al. [4] developed a web-based system that provides personalized advices based on the users' preferences for specific criteria, such as monetary costs, time costs, or the effect on their comfort, reporting promising results in first tests. In contrast to Ford et al.'s [4] conversational recommendation approach that provides recommendations only if users are searching for them, we aim at developing a system that is able to recommend energy-saving actions whenever users can save energy, without prior interaction. For example, the system could remind people of switching off the

light if they leave a room over a longer period of time, or it could recommend buying more energy-efficient devices if the users have the financial means.
Since such actions often entail crucial changes to users' habits, the system needs to provide certain sensitivity in persuading people to execute the recommended actions. Therefore, it could incorporate the findings of models on behavior change [3,10] (Section 2) into the systems decision making process. The main contribution of this paper is an investigation of the Energy Cultures Framework [11]. It describes criteria by which people's energy behavior can be characterized and thus identifies target points for actions by which people could be supported in enhancing their behavior on a long-term basis, see Section 3. In order to find out, whether the framework can be incorporated to enhance our recommender system's ability to choose suitable energy-saving actions, an online survey was conducted to gain insights into different Energy Cultures' preferences for energy-saving actions. The design and results of this online survey are presented in Section 4. The last section concludes the work and presents future work.

2 Behavior Change

According to *Fogg's Behavior Mode*l [3] behavior changes are affected by three factors: people's *motivation* and *ability* to execute a given task, and the *trigger* to take action. A persuasive recommender system therefore should only recommend actions that users are able and also willing to perform if they are motivated by the system. Thereby, even a person with a low motivation could be persuaded to apply an action if the trigger presents an action that can be executed easily enough. The right time to provide the trigger could be chosen by a context-aware system that proactively and immediately provides recommendations whenever users are able to make use of them.

In comparison the *Transtheoretical Model (TTM)* [10] describes behavior change in five stages: First people are unaware, and unwilling to change their behavior (*Pre-contemplation*). Then they acknowledge their problems, are open to information about the problem behavior, and intend to change their behavior in the future (*Contemplation*). *Preparation* is the stage in which people are ready to change their behavior in the near future and thus plan actions. Then they take *Actions* to change their behavior. Finally, they try to sustain changes and to prevent relapses (*Maintenance*).

In summary, a persuasive recommender system for energy-saving actions should only provide recommendations that fit the users' *motivation* or *stage of change* as well as their *abilities*. While less motivated people should be provided with actions, such as “Turn off devices instead of leaving them in

stand-by mode", more energy-aware people with the possibility to go to work by bike or by public transport could get the corresponding recommendation. In order to understand and analyze the origin of people`s energy behavior a system could incorporate the Energy Cultures Framework [11].

3 Energy Cultures Framework

According to the Energy Cultures Framework [11], people's energy behavior is influenced by three major factors: People`s *material culture* describes their properties and surroundings, such as their home`s insulation or the possession of different means of transportation. However, these properties are not only rated by their number, but mostly in terms of energy efficiency. The factor *norms* summarizes people's attitude towards the topic “energy saving" by considering whether energy saving is an important issue for them in general and by considering their specific opinions, e.g., an adequate room temperature. Finally, persons' executed *energy practices* describe their actual behavior, such as the current energy consumption, or heating habits.

The factors also influence each other. For example, a good insulation (Material Culture) leads to a decreased need for heating (Energy Practices) and the resulting savings could motivate people to save even more energy (Norms). The factors also can be influenced by external actions such as laws or, e.g., an increased price of electricity. However, it is ineffective to develop only uniform approaches or guidelines to foster changes of people's energy behavior, since attitudes and prerequisites inside a population are very heterogeneous [11]. We aim at tackling this issue by a personalized recommender system.

To summarize, incorporating the Energy Cultures Framework into a recommender system, based on Fogg's Behavior Model, seems to be a good approach to develop a persuasive system that fosters energy-efficient behavior. While knowledge about users' material culture and already executed energy practices could facilitate the choice of applicable, but also serendipitous recommendations, users' norms along with their already executed energy practices could be an indicator for their level of motivation and thus also the acceptable difficulty of recommended energy-saving actions.

4 Online Survey

In order to evaluate whether users' preferences concerning energy-saving actions depend on their *(cognitive) norms*, *material culture*, and their already performed *energy practices* we conducted an online survey.

4.1 Method

The first part of the study (22 questions and statements) was aimed to collect information by which the participants could be clustered regarding their energy-cultural background. These questions were based on a detailed description of characteristics of four major Energy Cultures that were identified by Lawson et al. [9]: Energy Economic, Energy Extravagant, Energy Efficient, and Energy Easy. The statements had to be rated on a 5-point Likert-Scale from “not at all" to “in any case". The participants' norms were assessed by their answers on statements, such as “I don't worry about energy-saving.", “Saving energy is important. However, so far I perform only a few energy-saving actions." or “I really watch my energy consumption." It was also considered whether participants save energy to protect the environment. To assess participants' material culture they had to rate their home's insulation from “very bad" to “very good". Additionally, they had to state the utilized type of heating system, the number of household appliances, the number of cars and motorized two-wheelers, and the number of bicycles. Afterwards a per capita value for the amount of possessed things was calculated and all facts were rated regarding their energy efficiency.

In the second part of the survey the participants were confronted with 21 energy-saving actions that were compiled based on tips on energy-saving provided by the WWF[1], and the German BMUB[2]. They included changes on the buildings, buying more energy-efficient devices as well as changes to the users' behavior, see Table 1. They differed, amongst others, in terms of temporal as well as financial expenditure, and eventual inconveniences for the users. The participants had to rate their interest for the specific actions as well as their willingness to execute them on a 5-point Likert-Scale from 1 = “not interested/willing at all" to 5 = “very interested/willing". To access the participants current energy practices an additional option was added, 6 = “already applied".

---

1 http://www.wwf.de/aktiv-werden/tipps-fuer-den-alltag/energie-spartipps/stromsparen/

2 http://www.bmub.bund.de/themen/klima-energie/energieeffizienz/foerdermittelberatung/energiespartipps/

4.2 Results

In total, 32 women and 58 men took part in the survey. They included very heterogeneous people with different demographic backgrounds: (A) Young (21-30), unmarried students, on lower incomes, living in a rented flat or apartment, or in shared flats in a middle or big city; (B) Middle aged (31-50), married, unmarried or divorced, partly with children, working in all kinds of profession, on average incomes, living either in owner-occupied houses in more rural areas or in rented flats or apartments in big cities; (C) Older than 51, married, children that have often already moved out, working in all kinds of profession, on average to high incomes, living in middle to smaller cities or in rural areas, mostly in houses, either owner occupied or rented.

(1) • Air rooms periodically with windows fully opened instead of continuously with tilted windows • Turn down the heating in unused rooms • Turn off lights in unused rooms (2) • Vent radiators regularly • Close shutters or curtains to keep the temperature constant • Install energy-efficient lamps • Line-dry laundry outside • Change the refrigerators temperature setting to 7° • Turn off devices instead of leaving them in stand-by mode • Go by bike or by foot more often	(3) • Install draught stops around doors and windows • Replace/upgrade windows • Install smart or programmable thermostats • Replace appliances by more energy-efficient ones • Pull the plug of unused appliances • Use public transport more often (4) • Install/upgrade wall insulation • Install heat-reflective mats behind radiators • Replace/upgrade your heating system • Buy a hybrid or electric car • Do car sharing

Table 1. Collection of energy-saving actions clustered by the survey's results

First, participants were clustered depending on their material culture, their norms, and their energy practices. For each factor we calculated the participants' proportion of positive answers in comparison to the total amount of questions considered for the specific factor. Proportions higher than 60%

were assigned a positive appraisal. Proportions between 40% and 60% were rated neutrally, and proportions less than 60% were rated negatively. The distributions of the ratings for the three factors describing Energy Cultures are shown in Figure 1.

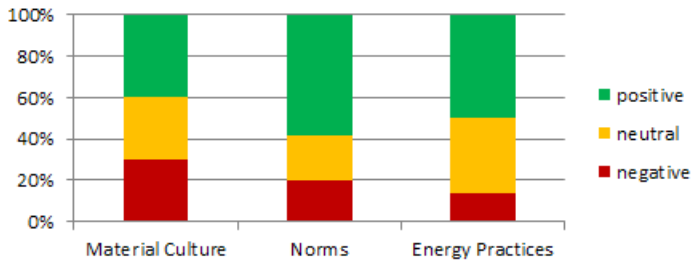


Figure 1. Assessment of the participants' material culture, norms, and energy practices

The actions could be categorized into four groups, see Table 1: (1) Well-known actions that most of the participants already are used to execute. (2) Actions that have already been applied by approx. half of the participants. Additionally, many of the remaining participants (40-70%) are interested and also willing to execute these actions. (3) Actions that approx. a third of the participants already perform, with approx. another third being interested in or even willing to perform them. (4) Actions that were executed by only a small number of participants (1-20%) and that also most of the participants were not interested in or willing to execute.

After categorizing participants as well as actions the provided mean ratings per category of participants were analyzed by a descriptive analysis to get a first impression of whether differing categories also had different preferences, see Figure 2.

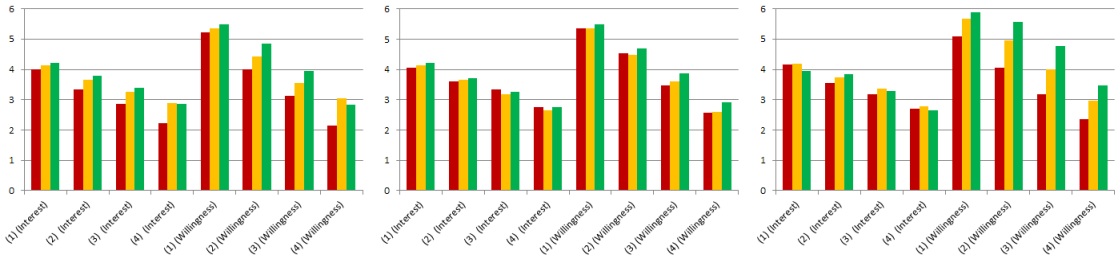


Figure 2. Mean ratings for actions of the categories (1) - (4) provided by members with a positive (green), neutral (yellow) or negative (red) appraisal for the investigated Energy Cultures factor: Norms (left), Material Culture (middle), Energy Practices (right)

The mean ratings of participants with differing norms showed differences in the groups' interest in specific energy-saving actions as well as their willingness to perform them. Participants with energy-friendly attitude in parts showed more interest and willingness than participants with negatively

assessed norms, especially concerning less popular energy-saving actions. Concerning the material culture, mean ratings of the participants showed only small differences. However, participants that already had a good material culture were more willing to apply the presented actions. Whether people already applied a lot of energy practices or not, had no effect on the mean ratings concerning their interest in the actions. However, since a positive appraisal concerning the criteria energy practices was based on a large number of already performed energy practices there was a major difference concerning the people's mean ratings related to their willingness to perform the recommended actions.

4.3 Discussion

A first analysis of the survey's results showed that people with a differing energy behavior have differing preferences for energy-saving actions. We think that by categorizing users into groups of people with a similar energy-cultural background, the system's decision-making process could be enhanced. For example, the survey's results let assume that users that already perform a lot of energy practices showed less interest in several of the presented actions because they already knew and applied them. Therefore, they should be provided with serendipitous recommendations. In contrast, people that perform only a few energy practices seemed to be not interested in several of the actions because they actually were not willing to perform them. This was especially apparent for actions that required higher monetary or timely costs, or that affected their comfort. They should be provided with actions that are easy to perform. In contrast, people with energy-friendly norms could be persuaded to apply actions that require more effort. Another advantage of incorporating the concept of Energy Cultures is that the system could be enabled to react to the users' behavior change by adapting the user models stepwise based on users' reactions to previous recommended actions.

5 Conclusion

We presented an idea of a recommender system that is aimed to foster energy efficient behavior by recommending personalized energy-saving actions. We aim at developing a system that provides only recommendations that fit the users' current state of change, motivation, and abilities. As a first step, we investigate whether the concept of Energy Cultures [11] could be a useful approach to model users' energy behavior including their motivation and

ability to apply different kinds of energy practices. A first analysis of an online survey's results showed promising results. In the near future the data gathered in the survey will be analyzed more in detail and a recommender system will be developed that utilizes the concept of Energy Cultures to provide feasible, useful and interesting recommendations to support people in enhancing their energy behavior.

Acknowledgements

This research is related to the IT4SE project, funded by BMBF. For more information visit http://www.it4se.net.

References

[1] R. M. Benders, R. Kok, H. C. Moll, G. Wiersma, and K. J. Noorman. New approaches for household energy conservation - in search of personal household energy budgets and energy reduction options. Energy Policy, 34(18):3612-3622, 2006.

[2] C. DiSalvo, P. Sengers, and H. Brynjarsdottir. Mapping the landscape of sustainable hci. In Proc. of the SIGCHI Conf. on Human Factors in Computing Systems, CHI '10, pages 1975-1984, New York, NY, USA, 2010. ACM.

[3] B. J. Fogg. A behavior model for persuasive design. In Proc. of the 4th Int. Conf. on Persuasive Technology, Persuasive '09, pages 40:1-40:7, New York, NY, USA, 2009. ACM.

[4] R. Ford, O. Sumavsk, A. Clarke, and P. Thorsnes. Personalized energy priorities: A user-centric application for energy advice. In A. Marcus, editor, Design, User Experience, and Usability. User Experience Design for Everyday Life Applications and Services, volume 8519 of LNCS, pages 542-553. Springer International Publishing, 2014.

[5] J. Froehlich, L. Findlater, and J. Landay. The design of eco-feedback technology. In Proc. of the SIGCHI Conf. on Human Factors in Computing Systems, CHI '10, pages 1999-2008, New York, NY, USA, 2010. ACM.

[6] L. Gamberini, A. Spagnolli, N. Corradi, G. Jacucci, G. Tusa, T. Mikkola, L. Zamboni, and E. Hoggan. Tailoring feedback to users' actions in a persuasive game for household electricity conservation. In M. Bang and E. Ragnemalm, editors, Persuasive Technology. Design for Health and Safety, volume 7284 of LNCS, pages 100-111. Springer Berlin Heidelberg, 2012.

[7] G. T. Gardner and P. C. Stern. The short list: The most effective actions U.S. households can take to curb climate change. Environment: Science and Policy for Sustainable Development, 50(5):12-25, 2008.

[8] C. Koehler, B. D. Ziebart, J. Mankoff, and A. K. Dey. Therml: Occupancy prediction for thermostat control. In Proc. of the 2013 ACM Int. Joint Conf. on Pervasive and Ubiquitous Computing, UbiComp '13, pages 103-112, New York, NY, USA, 2013. ACM.

[9] R. Lawson and J. Williams. Understanding energy cultures. In Conf. of the Australian and New Zealand Marketing Academy, pages 3-5, 2012.

[10] J. O. Prochaska and W. F. Velicer. The transtheoretical model of health behavior change. American Journal of Health Promotion, 12(1):38-48, 1997.

[11] J. Stephenson, B. Barton, G. Carrington, D. Gnoth, R. Lawson, and P. Thorsnes. Energy cultures: A framework for understanding energy behaviours. Energy Policy, 38(10):6120-6129, 2010.

Towards using Low-Cost Opportunistic Energy Sensing for Promoting Energy Conservation

Nuno J. Nunes, Lucas Pereira, Valentina Nisi
Madeira Interactive Technologies Institute
Funchal, Portugal
njn@uma.pt, lucas@m-iti.org, valentina@uma.pt

Abstract. This position paper discusses how to leverage low-cost energy sensing to opportunistically develop activity-based approaches to energy conservation. Based on our extensive experience developing low-cost sensing infrastructures and long-term deployment of ecofeedback systems, we discuss the possibility of unobtrusively inferring domestic activities from the overall aggregated energy consumption of households. We then postulate how the combination of this information with daily household activities could lead to more effective and meaningful ways to re-aggregate residential energy consumption for the purpose of ecofeedback. Here we briefly present a practical approach towards this new research direction that leverages HCI related methods, in particular using the day reconstruction method to provide semi-supervised approaches for automatic detection of household activities.

1 Introduction

Following the vision of the Internet of Things, household devices and appliances such as mobile phones, TVs, refrigerators, and even kettles and toothbrushes are increasingly endowed with more powerful computation, sensing, and communication capabilities. These "smart" devices can provide valuable information about the environment and the activities taking place around them. For instance electricity meters, environmental sensors, surveillance cameras and thermostats can be used to infer on-going households activities (e.g., cooking, cleaning, washing, leisure, etc.) and provide estimates of weekly and even daily occupancy of households.

In this position paper we describe our initial research results looking at how to leverage non-intrusive load monitoring (NILM) technology to opportunistically develop activity-based approaches to energy conservation. Here we combine sensing information from a NILM system, which includes energy consumption, energy events and user interaction events, with family routines obtained through daily reconstruction method. By sensing human activities and learning how these drive the use of energy in households, we

aim to facilitate future scenarios that could provide meaningful activity based feedback. We speculate that our approach could provide recommendations and/or automation possibilities, including optimizing micro-generation and off-the-grid household scenarios.

2 Related Work

In many domains (e.g., face recognition) the publicly availability of datasets was fundamental in improving machine learning and data mining techniques. Currently the fields of energy, environment and sustainability research are also seeing the emergence of publicly available datasets. Within these areas the Non-Intrusive Load Monitoring (NILM) community is particularly prominent given the need for extensive use of machine learning and data mining techniques.

Research in this field aims at disaggregating and estimating the consumption of individual appliances by means of applying machine learning techniques to the aggregated consumption signals [1]. NILM public datasets are expected to help researchers create more systematic evaluation processes that can be used across the different existing approaches. Moreover, as current research related to the human side of energy monitoring suggests, householders tend to associate their consumption with everyday activities (e.g. cooking, leisure, cleaning) [2]. We anticipate that future research will attempt to leverage the potential of existing datasets to automatically recognized and extract these activities.

Activity recognition is a long-established field of research. Previous work looked at human trajectories, interactions with objects or social activities [3]. However, with the exception of [4], most approaches neither target energy conservation, nor use the electricity consumption as an input variable for the recognition of human household activities.

3 Combining Consumption and User Activities

Our approach tries to tackle the known limitations of current ecofeedback systems, which focus on increasing efficiency by raising end user awareness of how their actions impact the use of energy. Our previous research showed that energy disaggregation strategies, commonly used in ecofeedback systems, are overwhelming for most users that lose interest and show relapsing behaviors in their energy conservation actions [5]. From the initial challenge of creating effective low-cost disaggregation strategies we faced the new

problem of generating meaningful strategies to re-aggregate consumption data that could effectively lead to long-term sustainable energy conservation practices in domestic environments.

3.1 Low-cost Non-intrusive Sensing of Consumption

NILM is considered a low cost alternative to attaching individual sensors on each appliance. Our research group developed a hardware and software platform [6] to enable the quick deployment of long and short-term studies of ecofeedback technology and at the same time serve as a research platform for developing NILM algorithms and techniques, as well as annotated public datasets [7]. The research involved several deployments of different ecofeedback systems, including both qualitative and quantitative evaluation of the user interactions. The overall goal was to raise the understanding and the awareness towards motivating people to consume more sustainably.

Our hardware/software platform evolved according to the requirements of the different deployments. From our initial setup was located in the mains and the ecofeedback was provided on-site (Figure 1 – left), it was later expanded into a full-fledge sensing platform with a dedicated multi-channel DAQ for multiple household studies of the long-term effects of ecofeedback (Figure 1 – right).

Figure 1. Low-cost NILM sensing system: single-house (left), multi-house (right)

3.2 Collecting Activities: Day Reconstruction Method

The Day Reconstruction Method (DRM) [8] is a well known survey method for characterizing daily life experiences. Users are asked to reconstruct their activities and experiences of the previous day without the burden of having to systematically use memory to remember past event, thus reducing the recall bias. The DRM can be thought of as a two step process: (i) users are asked to keep a diary of the previous day where they will list activities as being a sequence of episodes; (ii) every episode listed in the diary is then described by answering a series of questions about the situation and the feelings that they have experienced.

3.3 Research Questions

By combining low-costs NILM sensing with DRM we can focus on household activities to gain a better understanding of how and when energy is consumed and (micro) generated in a domestic environment. In this context the following research question emerges:

RQ1. Can we unobtrusively infer domestic activities from the overall aggregated energy consumption of the household?

Previous research in activity recognition typically uses data from a variety of sensors (e.g., presence, sound, cameras, etc.), which involves a dedicated infrastructure that in turn requires additional costs, energy consumption, management and maintenance. We instead intend to use an opportunistic approach based on our NILM infrastructure, weather and other environmental sensors and device-level consumption data and additional sensors for ground truth and calibration. We anticipate that in the future this data could be accessible by smart appliances. Therefore an additional research question emerges:

RQ2. Are domestic activities a more effective and meaningful way to re-aggregate residential energy consumption for the purpose of ecofeedback?

Regardless of the known shortcomings of ecofeedback systems [5], research demonstrated several important findings in terms of guidelines for information presentation. For instance it is known that traditional kWh representation (or even CO2) is not an adequate form of feedback in particular when presented in large aggregated monthly data. Conversely displaying appliance level consumption information is overwhelming and ineffective, as users tend to lose interest in detailed information relapsing to previous behaviors. With this research question we aim at investigating if re-aggregating consumption by domestic activity is more meaningful and effective as a basis for ecofeedback strategies and also for generating recommendation for energy conservation.

4 Exploratory Research

In order to test the feasibility of our approach we used our own NILM Dataset combined with data collected using the DRM in one of the actual deployments from our research.

4.1 SustData Dataset

The SustData dataset (freely available at http://aveiro.m-iti.org/data/sustdata) contains over 50 million individual records of electric energy related data, spanning a total of 1144 distinct days since July 2010 coming out of four distinct deployments of our sensing infrastructure. Currently the dataset contains over 11 million individual power events across all the four deployments. As stated in [7] a quick inspection of the results immediately reveals the high values for the daily standard deviation, which is a clear indicator of the large difference in the number of power events across the different houses, for instance, in the second deployment four houses have over 1000 daily power events on average, while seven houses have less than 300 power events per day. More detailed analysis (see Figure 2) provides a clear mapping between energy consumption and power events throughout different periods of the day and the week. We conducted additional analysis correlating family size and other independent variables leading to some interesting patterns among different families and households. For instance as reported in [5] there is a correlation between consumption, power events and the presence of children in the household, but not, for instance, between households with three or four people.

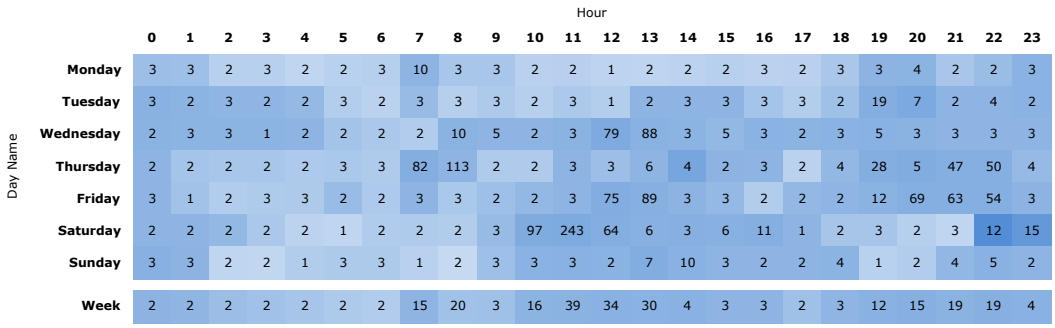


Figure 2. Average Consumption (color coded) and Power Events (nr.) for one Household (# 22) per Day of the Week and Hour of the Day

4.2 Family Activities

Datasets like SustData are becoming more popular and present an opportunity to explore opportunistic sensing of human activities. However collecting ground truth for human activities, in particular for long periods is difficult. Either the resident needs to keep record of all the activities, which is not convenient, or additional sensors are needed (for instance cameras or plug-level sensors) to label each activity, which is costly and not practical.

Role in family: daughter **Age:** 11 **Activity:** watch TV **Category:** entertainment **Start:** 12:30 **End:** 13:00 **Devices:** TV **Location:** living room **Mother Location:** outside of the house **Father Location:** kitchen **Other kid location:** bedroom	**Role in family:** mother **Age:** 33 **Activity:** clean kitchen **Category:** cleaning **Start:** 21:20 **End:** 21:35 **Devices:** lights kitchen, lights living room, TV **Location:** living room **Father Location:** living room **Kid location:** living room

Figure 3. Examples of Activities

The study reported in [9] consisted of two parts: a) a 1-day diary of all members of the family followed by b) interviews the day after. Using the Day Reconstruction Method (DRM) [6], we asked all family members to list, in a chronological order, the activities they performed while being within the house during the reported day. For each activity they provided a brief name and start and end time. Following the complete reconstruction of all daily activities, participants were asked to provide more detailed information for each activity (see Figure 3). This information was: a) electrical devices that were directly or indirectly used in the activity, b) all family members' locations (kitchen, living room, bedroom, outside house, other part of house).

Category	Activity (occurrences)	Appliances
Cleaning	Wash dishes (3) Dehumidify air (1)	Dishwasher; Dehumidifier
Entertainment	Play (8) Watch TV (22)	Unknown game console; Wii game console; PlayStation game console; Television; Personal Computer
Exercise	Workout (1)	Treadmill; Television
Laundry	Wash Clothes (2) Iron Clothes (3)	Washing Machine; Iron
Meal	Breakfast (13) Lunch (3) Dinner (12)	Microwave; Coffee grinder; Coffee machine; Kettle; Stove; Oven; Bottle heater; Fridge; Television; Personal computer; Toaster; Smoke extractor
Personal Care	Shower (4) Dry Hair (10)	Water heater; Smoke extractor; Fan
Work	Study (2) Work (13)	Personal computer; Laptop; Television

Table 1. Activities grouped by category

A total of 12 families participated in this study and reported on a total of 15 different activities that used 28 different appliances. These activities were later grouped into seven categories as shown in table 1, highlighting some interesting patterns in the family activities. For example, it becomes evident that some activities are much more common than others (watch TV was mentioned 22 times whereas working out and cleaning air were only mentioned once each). Likewise, it is also possible to see that some activity categories are much more prominent than others, particularly the

entertainment and having a meal categories that together account for roughly 60% of the mentioned activities and use about 57% of the listed appliances.

On average each house reported 12 activities (SD = 3,46), yet some houses were more active than others (e.g., house #22 reported 17 activities during the day, four houses reported nine activities, whereas on the lower end, house #2 reported only four activities). A more detailed analysis (see Figure 4) provides a clear mapping between energy (consumption and power events), user activities and appliances for two of the families that took part in the diary study. This analysis revealed some interesting usage patterns for some appliances, namely the TV that appears associated with several different activities (watch TV, play, leisure, lunch and dinner).

Likewise, this kind of analysis also suggests that families may not report all their daily activities. This is particularly evident in the house on the left side of Figure 4, which at 23:00 shows 40 power events that are not reflected in the reported activities (such a high number of power events would normally be associate with a cycling appliance like a clothes washer, which could indicate activities in the laundry category).

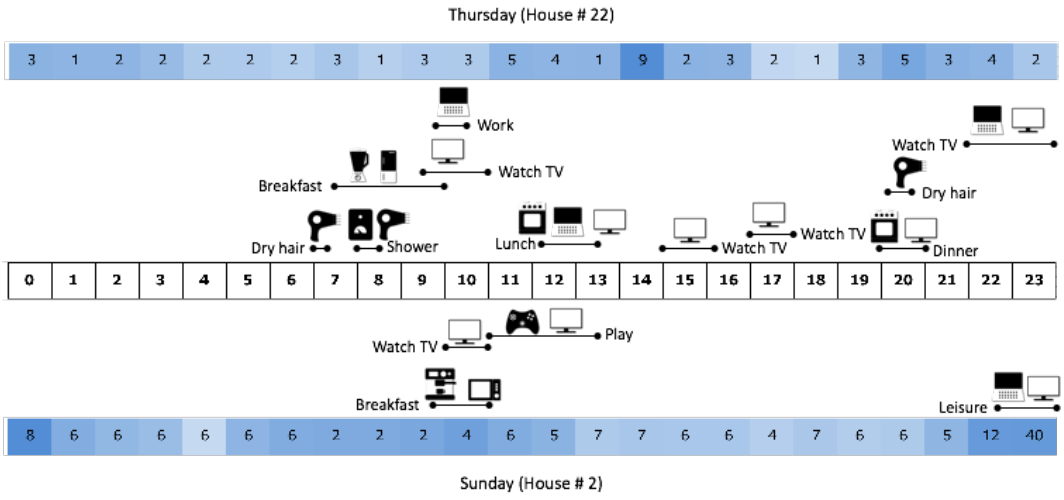


Figure 4: Consumption (color coded), Power Events (nr.), Activities and Appliances for two sample days of the week from two households in SustData

5 Discussion and Conclusion

The main output of this research will be an activity-based sensing approach to energy conservation in domestic environments. By combining NILM with activity based sensing we want to develop approaches to energy conservation that move beyond traditional ecofeedback systems and anticipate distributed micro-generation scenarios leading to important changes in energy sustainability and ultimately the utility business. To this end we anticipate to

provide a combination of i) actionable recommendations for energy conservation including those that take advantage of the availability of renewable sources, and ii) suggesting novel approaches for in-house automation that could leverage smart appliances and grid supply / demand balance. Our approach uses domestic activities and production/consumption parity to overcomes the known limitations of traditional ecofeedback systems.

In order to achieve these goals we aim at conduct a test-bed of 100 households to collect aggregated consumption data through our NILM framework, extended with micro-generation data, weather and other environmental parameters and finally appliance level ground truth data to generate semi-automatic labels. This data will come from diverse socio-economical backgrounds and will enhance our existing SustData public dataset [7] to provide the research community with a powerful tool to conduct additional research in the field of energy conservation. The final objective is to contribute to enhance the state of the art in ecofeedback technologies from opportunistic activity based sensing of domestic human activities related to energy.

References

[1] M. Zeifman and K. Roth, "Nonintrusive appliance load monitoring: Review and outlook," *IEEE Trans. Consum. Electron.*, vol. 57, no. 1, pp. 76–84, 2011.

[2] E. Costanza, S. D. Ramchurn, and N. R. Jennings, "Understanding Domestic Energy Consumption Through Interactive Visualisation: A Field Study," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, New York, NY, USA, 2012, pp. 216–225.

[3] J. K. Aggarwal and M. S. Ryoo, "Human Activity Analysis: A Review," *ACM Comput Surv*, vol. 43, no. 3, pp. 16:1–16:43, Apr. 2011.

[4] H.-Â. Cao, T. K. Wijaya, and K. Aberer, "Estimating Human Interactions with Electrical Appliances for Activity-based Energy Savings Recommendations: Poster Abstract," in *Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings*, New York, NY, USA, 2014, pp. 206–207.

[5] L. Pereira, F. Quintal, M. Barreto, and N. J. Nunes, "Understanding the Limitations of Eco-feedback: A One-Year Long-Term Study," in *Human-Computer Interaction and Knowledge Discovery in Complex, Unstructured, Big Data*, A. Holzinger and G. Pasi, Eds. Springer Berlin Heidelberg, 2013, pp. 237–255.

[6] L. Pereira, F. Quintal, N. Nunes, and M. Bergés, "The Design of a Hardware-software Platform for Long-term Energy Eco-feedback Research," in *Proceedings of the 4th ACM SIGCHI Symposium on Engineering Interactive Computing Systems*, Copenhagen, Denmark, 2012, pp. 221–230.

[7] L. Pereira, F. Quintal, R. Gonçalves, and N. J. Nunes, "SustData: A Public Dataset for ICT4S Electric Energy Research," presented at the International Conference on ICT for Sustainability, Stockholm, Sweeden, 2014.

[8] A. B. K. Daniel Kahneman, "A survey method for characterizing daily life experience: The day reconstruction method," *Science*, vol. 306, no. 5702, pp. 1776–80, 2005.

[9] M. Barreto, E. Karapanos, and N. Nunes, "Social Translucence as a Theoretical Framework for Sustainable HCI," in *Human-Computer Interaction – INTERACT 2011*, P. Campos, N. Graham, J. Jorge, N. Nunes, P. Palanque, and M. Winckler, Eds. Springer Berlin Heidelberg, 2011, pp. 195–203.

Private Focus Portals to Shared Energy Visualizations

Chi Tai Dang[1], Masood Masoodian[2], Elisabeth André[1]
[1] Human-Centered Multimedia Lab, Augsburg University, Germany
[2] Department of Computer Science, The University of Waikato, New Zealand
[dang, andre]@informatik.uni-augsburg.de
masood@waikato.ac.nz

Abstract. Large interactive displays provide potential platforms for collaborative visualizations to support groups of co-located people in interacting with shared information spaces. In these settings, it is often necessary to provide the individuals with their own private views of the shared visualizations. In this paper, we present a prototype system that allows users to get private views of their areas of interest (i.e. focus) within the larger shared visualizations (i.e. context) displayed on tabletop surfaces through portals provided on mobile devices. We demonstrate the potential of this system for visualization of collective and personal energy consumption data with the aim of supporting smart energy applications.

1 Introduction

Collaborative visualizations allow groups of people, with potentially different expertise, to combine their individual analytic power to tackle complex problems from different angles in concert or independently [11]. Various visualization systems have been proposed to support synchronous co-located collaboration [5]. These systems often use multiple displays, where a large shared display is used in combination with a number of smaller private displays (e.g. mobile devices or laptops), each belonging to one of the group members [12, 1, 8].

Collaborative visualizations need to support two different types of activities of group members, in working together (*coupled*) and working alone (*decoupled*) [9]. In their review of coupling in co-located collaborative visualizations, McGrath et al. [8] identify three problems with the use of a single tabletop display in these settings: 1) adding private views to the shared display uses valuable screen real-estate, 2) these views have to be managed as group members move around the table, and 3) such private views are always visible to others, reducing the degree of decoupling that can be achieved. Based on this observation McGrath et al. propose a *Branch-Explore-Merge* approach to supporting coupled and decoupled modes of interaction in co-located collaborative visualizations [8].

In this paper, we investigate this type of *coupled/decoupled* form of group interactions in co-located collaborative visualizations. We introduce our prototype system, called *Show-Me*, which allows individual users to focus on their own data and areas of interest using their private mobile devices, while sharing a common visualization context displayed on a public tabletop display. Our scenario of use is that of colleagues working in various offices in a building who view, share, and compare their personal energy usage data with those of others using a visualization designed for comparing temporal energy consumption data.

2 Energy Usage Visualization Scenario

As mentioned, co-located collaborative visualization tasks often involve coupled and decoupled activities with seamless transitions between them. In such visualizations, it is necessary to provide individual users with their own personal view of the public information space shown on the large shared display. We assume that the public visual space in its entirety acts as the context, while the views provided to individual users act as their own personal area of focus. This area of focus may have a different level of magnification to the context (visual zooming), or it might have a different level of information detail or content (semantic zooming)[1].

Our demonstrative scenario here is based on visualizations used for comparison of energy consumption data. We use a visualization called *time-pie* [6], designed to support comparison of personal and collective energy consumption data. Time-pie is an example of radial visualizations, which have been successfully used on mobile devices to motivate energy awareness [10]. Figure 1-left provides a sketch of the time-pie visualization, showing the amount of energy (in percentages) used by three types of devices during a 24-hour time period. The entire 24-hour period is divided into twelve 2-hour time slices, and the size of each slice is proportional to the amount of energy used during that time period in relation to the entire day. Similarly, within each slice, the amount of energy used by each device is represented as a percentage area of the entire slice. Time-pie also shows contextual information (e.g. outside temperature, the number of people in the office, etc. for each time slice) around its circumference.

We use this visualization to allow people working in different offices in a university building to compare the energy consumption of their own offices

---

1 For a detailed review of focus and context interfaces, see [2].

(or devices) against those of others. The time-pie visualization of the offices in a building is shown on a tabletop surface placed in a communal place in the building (e.g. coffee-room). People working in various offices in the building can view energy consumption of all the offices on the tabletop. Individuals may also wish to compare and share their own energy consumptions, in terms of their entire office or individual device types, with others. We consider this type of comparison to be a private activity, at least initially, or until the individuals are prepared to share their views with others. In this case, one could imagine individuals using their own mobile devices (e.g. tablets) as physical portals to view and focus on their own personal energy consumption data in the context of the collective energy usage visualization shown on the public tabletop display.

Multiple users could also share and compare visualizations of their personal usage data with each other, in the context of the visualization for all the offices or devices in the building. This would allow them, for instance, to discover patterns, identify interesting usage behavior, etc. in an attempt to collectively, as well as individually, make sense of their energy consumption data.

We consider the radial form of the time-pie visualization (and its Cartesian variation, called *time-stack* [7]) to be more suitable for a tabletop display, where different group members in a co-located collaborative setting will view the same visualization from different orientations. As Isenberg and Carpendale [4] point out (with reference to a study by Wigdor et al. [13]), for collaboration around tabletop displays care must be taken in choosing visualizations that are less susceptible to distortions due to changes of orientation.

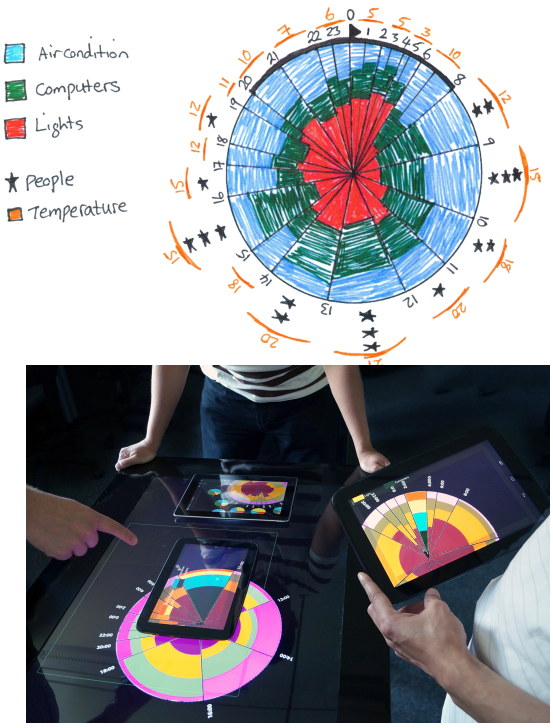


Figure 1. Sketch of the time-pie visualization (left), from [6]. A co-located collaborative visualization session using *Show-Me* (right).

3 Prototype System

The current *Show-Me* prototype utilizes a tabletop display to show the shared visualization of the context, while individual group members use their mobile tablet devices as portals to view visualizations of their own areas of interest (i.e. focus). Figure 1-right shows a co-located collaborative visualization session with three people using *Show-Me*. In this setting, various modes of coupled/decoupled group interactions are supported. These are:

1. *Coupled interaction around the tabletop without tablets* (Figure 2-left). In this mode the tabletop display is used to provide the shared visualization of the context, with the usual zooming, panning, and other forms of interaction possible.

2. *Decoupled interaction around the tabletop with tablets* (Figure 2-middle). In this mode the tabletop display provides the shared visualization context, while individual tablets provide private visualizations of the areas of focus. Group members may "hover and move" their tablets over the tabletop to get focus+context type visualizations (person *A* in Figure 2-middle). Alternatively, they may also move away their tablet from the tabletop (after the two have been

synchronized by simply placing the tablet on the tabletop), to get overview+detail type visualizations (person *B* in Figure 2-middle).

3. *Coupled interaction around the tabletop with tablets* (Figure 2-right). In this mode the public visualization is shown on the tabletop display, while individual tablets are used to share private visualizations of the individual areas of focus. The individual group members may "place and drag" their tablets on the tabletop display to share a focus+context visualization of their private views over the public view.

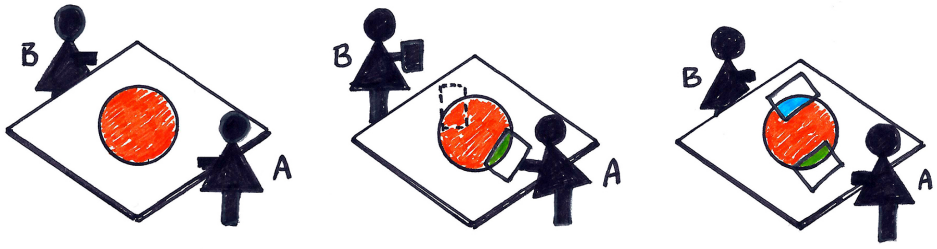


Figure 2. Coupled interaction around the tabletop without tablets (left), decoupled interaction with tablets (middle), and coupled interaction with tablets (right).

Although in this section we present these three coupled/decoupled modes of interaction separately, other combinations of them are also possible. For instance, in a modified coupled variation of the decoupled mode (Figure 2-middle) the group members may show each other their private visualizations using their tablets in the context of visualization shown on the tabletop.

It should also be noted that the type of focus provided on the individual tablets, in both coupled and decoupled modes, in most cases is likely to be a semantic focus (e.g. using semantic zoom). This would allow viewing private visualizations at various levels of semantic detail (as well as visual detail) over the shared public visualization.

3.1 Coupled Interaction around the tabletop without tablets

In this coupled mode of group interaction two or more people would use a shared visualization of collective energy consumption data shown on the tabletop display to view and compare energy use by different types of devices (Figure 3-left) or different offices (Figure 3-right) in the building. *Show-Me* allows selection of different visualizations of public data sets, as well as allowing users to zoom, select, and interact with different sections of the shared visualization.

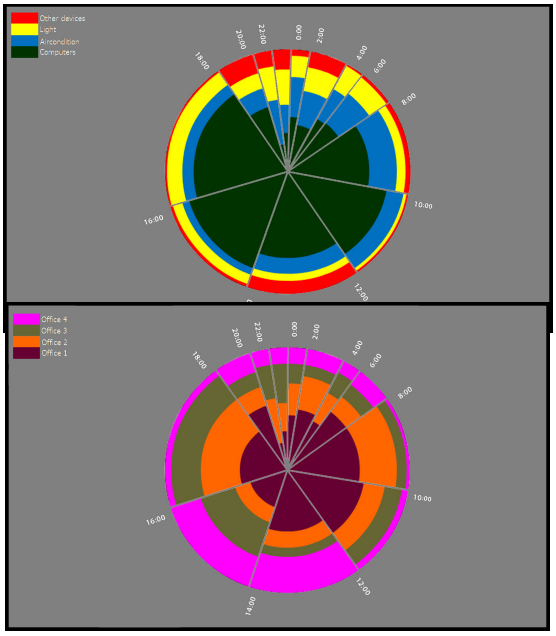


Figure 3. Public time-pie visualization of energy consumption by all the devices (left) and by all the offices (right).

3.2 Decoupled Interaction around the tabletop with tablets

In this decoupled mode of co-located interaction different group members will use their mobile devices (tablets in this case) to view and interact with visualizations of their personal (or even collective) energy consumption data. The visualization shown on the tablet will be considered the focus of the user, while the visualization shown on the tabletop will act as the context.

Once a tablet has been synchronized with the tabletop, it can then be used as a physical private focus portal to the public visualization shown on the tabletop display. As mentioned earlier, the focus can be changed by panning and zooming visually (i.e. change of magnification), or semantically by overlaying other visualizations or changing levels of information detail. Each tablet (and its user) are associated with a particular office, and as such, they are allowed to only view their own personal usage data in detail, or use other visualizations on their assigned tablets.

Also note that users can put their tablets on the tabletop, or lift, hover, and move them above the tabletop, to get their own focus+context views. Lifting the tablet completely and moving it away from the tabletop allows keeping its view in synch with the tabletop, and provides an overview+detail mode of operation.

3.3 Coupled Interaction around the tabletop with tablets

Coupled interaction around the tabletop could also include the use of tablets to share details of personal energy consumption, which are not available through the collective visualization of the tabletop display. For instance, individuals may wish to view and share visualizations of their personal energy consumption data with the aim of comparing and identifying interesting patterns of use.

Figure 4-left shows one such case, in which two people from offices 1 and 4 are using their tablets to compare their private energy consumptions by different devices between 8:00 and 10:00 in Office 1 with those of Office 4 between 12:00 and 14:00.

Note that although in the current version of our prototype it is not possible to layer tablets on top of one another to combine their areas of focus through physical layers, this is something that we are intending to investigate in the future. For now, if two tablets need to focus on different parts of the same area of the shared tabletop visualization (e.g. offices 1 and 4 between 12:00 and 14:00), then one or both of the users need to lift their tablets off the tabletop and interact with the tablet surface to move their base of focus to the same region.

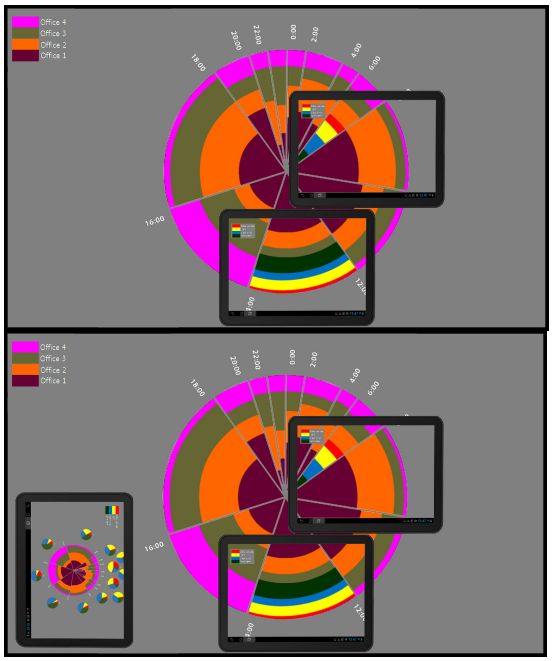


Figure 4. Public time-pie visualization of energy consumption by all the offices in coupled interaction around the tabletop with tablets. The left picture shows details of devices for Office 1 between 8:00 and 10:00, and Office 4 between 12:00 and 14:00. The right picture shows a combination of tablets, each with a different type of details for one of 3 of the offices.

Clearly it is not necessary to have the same type of focus visualization (e.g. semantic zoom) in all the tablets used in this form of coupled interaction. Figure 4-right shows an example of a case where the tablet on the left is providing a different visualization of the details of energy consumption by all the devices in one of the offices for the entire day.

4 System Implementation

The current *Show-Me* prototype has been designed as a distributed multi-surface (multi-display) system, consisting of application components for the Microsoft PixelSense tabletop, and Google Android-based tablets. The tabletop application was developed using the PixelSense Surface SDK (WPF) and shows the shared public visualization on the tabletop display surface. Tablet application, on the other hand, was developed using the Java programming language for the Android platform and provides the private focus views.

Both applications utilize a multi-platform MSE (Multi-Surface Environment) framework, called *Environs* [3], as a software layer integrated into the *Show-Me* applications for each participating device. By means of this framework, the devices span a so-called application environment, which is similar to a Peer-to-Peer network wherein the devices are loosely coupled together and communicate with each other directly.

To create and show the private focus views on the tablets, *Show-Me* leverages interactive video portals between devices provided by the *Environs* framework. Those video portals can be overlaid, with multiple static or dynamically changing content, before streaming to destination devices. Our prototype makes use of this feature to provide additional private visualizations (e.g. usage data for different offices) to individual users in coupled and decoupled interaction modes.

In order to track the mobile devices, our prototype makes use of the Byte tags[2], which are natively supported by the PixelSense tabletop.

5 Conclusions

Here, we have presented our prototype *Show-Me* system which utilizes a shared tabletop display with a combination of mobile tablet devices, to create a multi-surface environment for comparing and analyzing personal and

---

2 http://www.microsoft.com/download/en/details.aspx?id=11029

collective energy consumption data using the time-pie visualization. We intend to evaluate the effectiveness of *Show-Me* in supporting our proposed co-located collaborative energy usage visualization and analysis tasks. We will then modify and extend our system based on the results of this study.

Acknowledgements

This work has been supported by the IT4SE project (for details, see http://www.it4se.net).

References

[1] Cheng, K., Li, J., and Mueller-Tomfelde, C. Supporting interaction and collaboration on large displays using tablet devices. AVI '12, ACM, 774–775.

[2] Cockburn, A., Karlson, A., and Bederson, B. B. A review of overview+detail, zooming, and focus+context interfaces. ACM Computing Surveys 41, 1 (Jan. 2008), 2:1–2:31.

[3] Dang, C. T., and André, E. A framework for the development of multi-display environment applications supporting interactive real-time portals. EICS '14, ACM, 45–54.

[4] Isenberg, P., and Carpendale, S. Interactive tree comparison for co-located collaborative information visualization. *IEEE Transactions on Visualization and Computer Graphics 13*, 6 (2007), 1232–1239.

[5] Isenberg, P., Elmqvist, N., Scholtz, J., Cernea, D., Ma, K.-L., and Hagen, H. Collaborative visualization: Definition, challenges, and research agenda. *Information Visualization 10*, 4 (2011), 310–326.

[6] Masoodian, M., Endrass, B., Bühling, R., Ermolin, P., and André, E. Time-pie visualization: Providing contextual information for energy consumption data. IV '13, IEEE Computer Society (2013), 102–107.

[7] Masoodian, M., Endrass, B., Bühling, R., and André, E. Visualization support for comparing energy consumption data. IV '15, IEEE Computer Society (2015).

[8] McGrath, W., Bowman, B., McCallum, D., Hincapié-Ramos, J. D., Elmqvist, N., and Irani, P. Branch-explore-merge: Facilitating real-time revision control in collaborative visual exploration, ITS '12, ACM, 235–244.

[9] Schmidt, K., and Simone, C. Coordination mechanisms: Towards a conceptual foundation of CSCW systems design. *Computer Supported Cooperative Work (CSCW) 5*, 2 (1996), 155–200.

[10] Spagnolli, A., Corradi, N., Gamberini, L., Hoggan, E., Jacucci, G., Katzeff, C., Broms, L., and Jonsson, L. Eco-feedback on the go: Motivating energy awareness. Computer 44, 5 (May 2011), 38–45.

[11] Tobiasz, M., Isenberg, P., and Carpendale, S. Lark: Coordinating co-located collaboration with information visualization. *IEEE Transactions on Visualization and Computer Graphics 15*, 6 (2009), 1065–1072.

[12] Voida, S., Tobiasz, M., Stromer, J., Isenberg, P., and Carpendale, S. Getting practical with interactive tabletop displays: Designing for dense data, "fat fingers," diverse interactions, and face-to-face collaboration. ITS '09, ACM, 109–116.

[13] Wigdor, D., Jiang, H., Forlines, C., Borkin, M., and Shen, C. Wespace: The design development and deployment of a walk-up and share multi-surface visual collaboration system. CHI '09, ACM, 1237–1246.

Presentation Methods to Inform Decisions about Energy Usage

Chris Killeen
University of Brighton
c.killeen2@uni.brighton.ac.uk

Abstract. This paper reviews some of the body of literature on the effects of feedback and goal setting in promoting energy conserving behaviour. One of several devices that are being installed into households across the UK, the E-on Ecometer, is examined and the design of its interface critically evaluated. This is done by examining the information presented by the device in the light of the finding in the previously examined literature.

1 Introduction

Following the 2008 Climate Change Act being passed into law, the UK government has committed to reducing the UK's carbon emissions by 80% by the year 2050[1]. As the domestic market accounts for 15% of the UK's greenhouse gas emissions [2], and 30% of the UK's energy use [3], the government has created a programme with the aim of installing a smart meter and an energy monitor into every household in the United Kingdom by 2020 [4]. The hope of this project is to reduce the energy demand of the domestic market and therefore reduce the carbon emissions required in generating energy for this sector [5].

Energy is not something that is often thought about in the household; a user simply plugs a device into a wall socket or turns on the central heating and the device works. As a result, energy has effectively become invisible to householders [3].

Smart metering was chosen as the method for driving this reduction in usage, along with other measures such as housing insulation [5] as there is evidence to suggest that by making energy usage visible to the consumer, they are better equipped to make decisions about the amount of energy they use [7].

With the projected costs of the smart meter roll out being in the region of £11 billion [8], it would be reasonable to ensure that the devices being installed into every home in the UK are capable of performing the task they were designed to do.

2 Existing studies – The effectiveness of feedback

Energy monitors give the user direct feedback i.e. the user receives instant usage information direct from their meter, in addition to indirect feedback on their bills as processed by their utility company [9]. There have been several studies into the effects of direct feedback in helping users reduce their energy consumption undertaken in recent years.

Feedback is defined by the Oxford English Dictionary in Darby [9] as"…Information about the result of a process or action that can be used in modification or control of a process or system … especially by noting the difference between a desired and an actual result."

There have been several studies into the effects of feedback in driving energy conservation behaviour. The results have been largely positive. In Faruqui, Sergici & Sharif's 2010 review of the pilot programmes of energy monitors around the globe, they found that although one study showed a negligible impact on energy use reduction, energy savings could be as high as 18% with an average energy saving of 7% [10].

Although the results of these pilots are promising, many of the pilots were short term, or combined with other energy reduction measures, such as time variable tariff rates or use of prepayment meters. This clouds the studies when it comes to measuring the benefits of feedback from energy meters.

Carroll, Lyons & Denny's [11] study into the effectiveness of energy monitors in increasing energy reduction knowledge found that whilst energy monitoring does have an effect in reducing usage, the effect is much greater when combined with variable tariff rates.

McCalley and Midden [12] discuss the value of goal setting alongside feedback when encouraging energy reduction. Their experiment involved integrated feedback within an individual device as opposed to feedback from a smart monitor but the behaviour being studied is comparable. The study found that if users set their own energy saving goals, rather than having goals imposed on them, they were more likely to save energy. This experiment also took personality type into consideration, breaking the sample group into "pro-social", who are more motivated by altruism than the "pro-self" group who are more motivated by self-interest.

The experiment was not without its faults, the experiment being carried out in a laboratory setting during one day, rather than the home environment where the users were less likely to change their behaviour as a result of being observed.

The positive effect of goal setting was supported by Abrahamse, Steg, Vlek, & Rothengatter in their 2007 study into the effect of goal setting and feedback

on energy usage [2]. They found that when set a 5% reduction target using a website rather than an energy monitor to give the feedback, users with feedback and goal setting had achieved a 5.1% reduction in usage, whereas the control group with no goal setting had not achieved this reduction.

Despite the positive results being shown by many of these short term studies, there has been very little research done into the long term effects of energy monitoring systems. Hargreaves, Nye and Burgess [13] have produced the best example of the very few studies looking at energy monitoring systems more than 6 months after installation. Their study involved interviews with 11 out of a group of 275 householders who had had an energy monitoring system installed for 12 months. The households had been equipped with smart meters of varying degrees of sophistication and feedback. The study found that energy monitors became "part of the furniture" after the novelty value of the meter had worn off. They also found that after an initial burst of energy saving behaviour when the monitors were installed, users settled on what they considered a "normal" level of usage and were content to try to stick to that level rather than to reduce it [13].

This idea that users may stick to "normal" levels of usage over time is echoed by Strengers' study that found that if consumption did not "…scream red" then this behaviour was seen as legitimate by users of energy monitors. [14]

2.1 Normative feedback

Possibly one of the most influential studies into using feedback to increase energy conserving behaviour was carried out by Cialdini and Schultz in 2004, looking at different ways of getting people in San Marcos, California to use lower energy fans as opposed to air conditioning (AC) to cool their homes. [15]

The experiment consisted of randomly selecting candidates and leaving different door hangers containing energy saving messages on the handles of the homes in the study. These messages included telling the householder how much money they could save, asking the householder to think about the environment, telling the householder how much energy their neighbours were saving and a control group with no messaging.

All the groups either increased their energy use in the trial period or had a very small decline that was not statistically significant, with the exception of the group who were told that their neighbours were using fans instead of AC. These households saw a drop of their average daily usage by 1.2 kWh per day.

This experiment was so successful that it inspired the creation of Opower, a company that works with utility companies to tell their users how their energy usage compares to similar homes, which in 2013 saw a 2tWh reduction in usage from households receiving information from Opower [16].

This method of promoting behavioural change has not yet been applied to energy monitors, although it is being used by energy companies, for example, E-on and First Utility, in the customer account sections of their websites [17.18].

2.2 Social feedback

André, Bühling, Endrass and Masoodian's paper [19] describes the two key types of feedback often addressed in displaying energy usage via an energy monitor; Normative and historical feedback. Historical feedback is where users can see their own usage over time in comparison to their current usage, allowing the user to better understand their consumption. [19]

This paper also goes on to identify social feedback as an area that should be looked at when addressing energy use reduction. This differs from normative feedback as it allows users to compare their usage with other individuals or groups, for example other individuals in a house or other groups within an office [19].

3 Visualisation of energy usage via energy monitors

The interface of the energy monitor is the most important part of the device, as this is where the information about usage is displayed to the user. There have been studies carried out into optimising the design of the interface. Börner, Kalz, Ternier and Specht's study into the user of visual interfaces to promote energy conservation at a university campus looked at different approaches that could be taken to encourage energy use reduction in a workplace environment. The study found that different interfaces were more effective at giving different types of feedback, with public displays to pass on information, individual displays to return personalised feedback and an element of gamification to give incentives to reduce energy usage. [20]

A more detailed study of the visualization techniques that could be used in energy monitors was carried out by Thomas Rist [21]. This study describes the objectives of visualizing energy data in the context of energy conservation and describes some of the common methods of visualizing this data. The paper discusses the benefits of displaying energy data in charts and diagrams, energy gauges and eco-visualization and ambient feedback, where for example, an

image of a garden is shown, with the garden looking healthy when the user is using less than the average energy to the garden looking dead and withered when the user is using more than average [21]. The paper also looks at Interactive visual data exploration, mixed media feedback and energy related games.

4 Case study: The E-on Ecometer

The E-on Ecometer was chosen as it is the only device accessible at the time of the study. It is an older model of the installed devices containing a monochrome LCD screen which displays the feedback, a series of labelled buttons to its right and a 4 smaller buttons below. It was chosen for the case study ahead of more state of the art energy monitoring systems, such as the USEM system [22] as it is an example of what has been installed recently as part of the smart meter roll out.

The user selects which information to view by tapping the buttons to cycle through the information displayed on the screen. Using the menu button below the screen it is possible to set an alarm to sound if the household uses more energy than a specified level. On the bottom of the device are a bank of lights coloured green, orange, red and blue. The green light shows when the household is using less than 150w of power, the orange shows when the household is using between 150w and 1kw, and the red shows if the household is using more than 1kw. The blue light flashes every minute [23].

The buttons to the right of the screen have both picture symbols as well as text labels to describe their functionality. The top button is labelled “Select”. This button allows the user to cycle between displaying energy usage information for electricity or gas.

The button below is labelled “Reading/Costs”. This allows the user to cycle between seeing their usage data in kWhs or in a monetary value, in this case pounds and pence.

The next button is labelled “Usage”. This button allows the user to cycle through their historical usage for gas or electricity, depending on which has been selected using the top button. The usage displayed can be cycled through to report daily, weekly, monthly or yearly usage.

The final button on the right hand side can be used to show the household’s emission of Kg of CO2 so the user can see the environmental impact of their usage. It is labelled “Emissions”.

Below the screen are 4 buttons that allow the user to access and navigate through the menu system. The menu allows users to change the brightness of the lights on the front of the Ecometer, and to set an alarm that will sound if

the cost of daily energy usage goes above an amount that can be determined by the user.

The LCD display shows a variety of different information depending on what the user has selected to view. The display will show what energy they are using at the current time as default, along with a bar graph showing historic consumption for the last hour. On pressing a button the screen becomes backlit to allow viewing in dim light.

The choice of a bar graph is a good one for displaying quantative historical data over time in a time series [24], showing the data in a way that is easy to process and pull relationship information from [25]. The major issue with the bar graph, and all the other bar graphs shown on the device, is that they do not show a scale, which means the user can't interpret the information the graph shows except for getting very general trend information.

The user can choose to look at their usage, in kWh, pounds and pence or in Kg/CO2 over an hour as described above, over a day, a week, a 28 day period or over a year, with the usage shown in minutes, hours, days or months respectively, although this is not clearly labelled on the graph. The figure shown above all the graphs will show the usage over that entire time period. This again gives the user no specific information about their usage, such as variations that would be seen between weekend and weekday, or between different seasons with any degree of clarity other than broad trends.

The user cannot for example, see their usage for a specific day and see exactly how much greater or lesser their usage was compared to other days. This problem with the unlabelled graphs is particularly apparent in the case of high usage households as the graphs appear to have an upper limit; different usage over an undisclosed threshold will appear to be the same (Figure 1). This poor information design means feedback about long term energy usage in high usage households is not available. The lack of clear visibility of past usage also makes it difficult to use the alarm feature as the user cannot easily work out what they are using on average. This makes setting energy saving goals arbitrary, which in turn will make them difficult to stick to so ultimately the user will not sustain motivation to reduce their usage [12]. This is particularly disappointing as it would be logical to assume that high usage households would be the primary target to reduce CO2 emissions.

Much information is available to the user through the E-on Ecometer, but because of the choices made in design of the interface, the feedback the device returns is of far less value than it could have been. This will make goal-setting extremely difficult thereby reducing influence on user's behaviour to use less energy.

There is also no normative feedback which has been shown to be very effective in stimulating energy reduction behaviour, allowing the user to see a socially derived baseline to compare their own usage to [15].

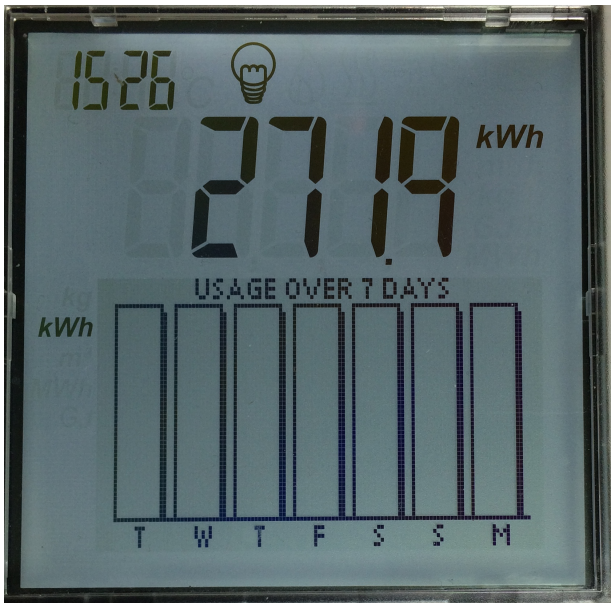


Figure 1. Yearly electricity usage display for a high usage household. (Author, 2014)

5 Conclusions

It is not obvious how the design of the E-on Ecometer interface aligns with the behavior changing objective motivating its introduction. The axis of displayed graphs is not labelled. The aggregated numbers it displays give a limited sub-set of potentially available usage information, reducing its value as feedback and therefore it's utility for self-setting usage reduction targets. This lack of valuable clear information may add to the backgrounding effect the device will suffer from, further reducing its effectiveness.

Considering the huge sum of money being spend on the smart meter rollout through either direct government funding or additional cost on the user's energy bill, it is surprising so few user studies have been carried out [19] to determine how effective the current interfaces are and what can be done to improve them.

The smart meter rollout is happening now, with energy monitors being installed into households throughout the United Kingdom every day. Validating guidelines for incorporating what is known about influencing behavior change into the design of such devices is urgently needed. This will require studies that monitor effectiveness over a period longer than a few months. This work needs to be undertaken as soon as possible as the rate of

installations is set to increase and without validation of the designs of the monitor interfaces, the project could easily turn into a very expensive waste of effort.

References

[1] Climate Change Act. Section 1, Elizabeth II. Chapter 27, 2008.
[2] Abrahamse, W., Steg, L., Vlek, C., & Rothengatter, T. The effect of tailored information, goal setting, and tailored feedback on household energy use, energy-related behaviors, and behavioral antecedents. Journal of Environmental Psychology, 27(4), 265-276, 2007.
[3] Burgess, J., & Nye, M.. Re-materialising energy use through transparent monitoring systems. Energy Policy, 36(12), 4454-4459, 2008
[4] Smart Energy GB. It's time to get Gaz and Lecky under control. [ONLINE] Available at: http://www.smartenergygb.org/ Last Accessed on 05/06/2015.
[5] Smart Energy GB. A Smarter Future. [ONLINE] Available at: http://www.smartener-gygb.org/what-are-smart-meters/smarter-future Last Accessed on 05/06/20145.
[6] Energy Saving Trust. Home Insulation. [ONLINE] Available at: http://www.energysavingtrust.org.uk/domestic/content/home-insulation Last Accessed on 07/06/2015.
[7] Schwartz, T., Stevens, G., Ramirez, L., & Wulf, V. Uncovering of making energy consumption accountable: A phenomenological inquiry. ACM Transactions on Computer-Human Interaction (TOCHI), 20(2), 12, 2013
[8] Smithers, R. Public may end up paying for obsolete smart meters that save little, MPs warn. The Guardian, [ONLINE] Available at: http://www.theguardian.com/environment/2014/sep/10/public-may-pay-obsolete-smart-gas-electricity-meters-save-little-mps-warn Last Accessed on 07/06/2015.
[9] Darby, S. The effectiveness of feedback on energy consumption. A Review for DEFRA of the Literature on Metering, Billing and direct Displays,486, 2006.
[10] Faruqui, A., Sergici & S. Sharif, A. The impact of informational feedback on energy consumption – A survey of the experimental evidence. Energy. 35 (4),1598-1608, 2010.
[11] Carroll, J., Lyons, S., & Denny, E. Reducing household electricity demand through smart metering: The role of improved information about energy saving. Energy Economics, 45, 234-243, 2014.
[12] McCalley, L. T., & Midden, C. J. Energy conservation through product-integrated feedback: The roles of goal-setting and social orientation. Journal of economic psychology, 23(5), 589-603, 2002.

[13] Hargreaves, T., Nye, M., & Burgess, J. Keeping energy visible? Exploring how householders interact with feedback from smart energy monitors in the longer term. Energy Policy, 52, 126-134, 2013.
[14] Strengers, Y. "Designing eco-feedback systems for everyday life." In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 2135-2144. ACM, 2011, 2011.
[15] Cialdini, R., & Schultz, W. Understanding and motivating energy conservation via social norms. William and Flora Hewlett Foundation, 2004.
[16] Lanskey, A. Alex Laskey: How behavioral science can lower your energy bill [Video file] Retrieved from: http://www.ted.com/talks/alex_laskey_how_behavioral_science_can_lower_your_energy_bill?language=en 2013.
[17] Opower. First Utility partners with Opower to help UK households cut their energy bills [Press release] Retrieved from http://www.first-utility.com/wp-content/uploads/oPower%20release.06.07.1L.pdf, 2011.
[18] Opower. E.ON launches online tool to help households understand and transform their energy use [Press release] Retrieved from http://opower.com/company/news-press/press_releases/84 2013.
[19] André, Elisabeth, René Bühling, Birgit Endrass, and Masood Masoodian. "Evaluating the Effectiveness of Visualizations for Comparing Energy Usage Data." FSEA 2014 2014: 5.
[20] Börner, Dirk, Marco Kalz, Stefaan Ternier, and Marcus Specht. "Pervasive Visual Interfaces to Change Energy Consumption Behaviour at the Workplace." 2014.
[21] Rist, Thomas. "Towards a More Responsible Use of Energy through Visualization of Energy Data." FSEA 2014: 9, 2014.
[22] Masoodian, Masood, Elisabeth André, Michael Kugler, Florian Reinhart, Bill Rogers, and Kevin Schlieper. "USEM: A ubiquitous smart energy management system for residential homes." (2014).
[23] E-on (n.d) Ecometer Guide [ONLINE] Available at: https://www.eonenergy.com/~/media/PDFs/For-your-home/saving-energy/Ecometer%20SED%20guide.PDF Last Accessed on 09/05/2015.
[24] Few, S. (2004a). Eenie, Meenie, Minie, Moe: Selecting the right graph for your message [ONLINE] Accessed at: http://www.perceptualedge.com/articles/ie/the_right_graph.pdf. Last Accessed on 08/06/2015.
[25] Few, S. (2004b). Tapping the power of visual perception [ONLINE] Available at: http://www.perceptualedge.com/articles/ie/visual_perception.pdf Last Accessed on 08/06/2015.

Watt-I-See: Probing Future Distributed Energy Scenarios

Clinton Jorge, Filipe Quintal, Nuno J. Nunes, Valentina Nisi
Madeira Interactive Technologies Institute
Funchal, Portugal
{filipe.quintal, clinton.jorge}@m-iti.org,
njn@uma.pt, valentina.nisi@gmail.com

Abstract. This position paper discusses a novel perspective on eco-feedback that goes beyond changing consumption behaviours through accountability. Here we describe an interactive installation that displays the ratio of current power generation sources and the percentage of renewables in the grid in an attempt to incite intrinsic values in order to persuade behaviour change. We build upon design insights that strike the complex balance between the awareness of the realities and complexities of user lifestyles on one hand, and a desire to influence their use and consumption behaviours and practices on the other hand. Our preliminary results show an increase in energy literacy and awareness as well as identify high consumer preferences towards simple, representative interfaces and ubiquitous immediate energy production feedback. Our study shows potential in terms of future scenarios for eco-feedback in distributed energy, micro-generation and other inevitable disruptive changes.

1 Introduction

Electricity is the cornerstone of our modern lives. Electrical devices support most of our daily activities as they power our social, governmental and health services. However, our society's increasing demand for electricity (and energy in general) is plainly unsustainable. Notwithstanding, disruptive changes seem inevitable due to a convergence of factors, including climate action, economic downturn, falling costs of energy resources and public policy incentivizing the adoption of new energy technologies. These factors are making distributed micro generation desirable and affordable for consumers. Recently, innovative companies such as Tesla, are pushing technology aiming at reducing the costs of batteries that could take houses with renewable micro-generation off the grid. Still, the increased penetration of renewables poses many challenges for energy management systems and does not necessarily generate a more sustainable future.

Unfortunately, fuel based production still accounts for approximately 40% of the worldwide total energy production of which 28% is accountable for residential use, with an estimation to increase to 32% by 2040 [10]. The

attempts to reduce residential consumption through increased efficiency are not reverting this tendency, as households now own more appliances than in the past. Small appliances proliferating in houses are currently estimated to account for half of the consumption, providing a significant margin for individuals to manage residential consumption. HCI research is looking at these phenomena typically from the perspective of encouraging sustainable energy consumption from end-users. Regrettably the results are far from the initial expectations [1,9].

This position paper follows a conviction in which "in order to assess the potential and effectiveness of HCI in environmental practice, it is necessary to inquiry into the contexts in which those practices arise, and to recognize the potential contradictions between the goals of our intervention and the forces that shape their deployments" [2].

2 Watt-I-See: Visualising Energy

HCI research mostly focuses on raising awareness and promoting sustainable energy consumption. These approaches usually rely on some form of eco-feedback technology, i.e. technology that provides feedback on individual or group behaviours with the goal of reducing environmental impact [4]. Our approach builds on the state of the art in a twofold way. First it explores how HCI and design research envisioned energy and, in particular, the tension between the seamless and ubiquitous nature of energy as a service provided to consumers, and the inherent intangible and invisible nature of the underlying commodity [6]. Secondly it tries to go beyond traditional eco-feedback by combining production and consumption information coming from a medium size closed grid with a high penetration of renewables.

Both these approaches have been tested in small isolated studies in which the source of the energy was included in the eco-feedback visualization (e.g. [3,7]), and in studies that aimed to present the intangibility of energy (e.g [5]).These studies reported an increase in consumer's interest and knowledge about electricity consumption.

Watt-I-See (WISE) is an interactive installation constructed with the ambition to raise awareness about energy production. Our goal with WISE was to first explore a tangible design that could reconcile the seamless and ubiquitous nature of the electricity. Secondly to approach the design challenge and go beyond the conventional micro-generation scenarios, combining production and consumption information from a medium-size grid setup. The aim with WISE was to provide actionable design guidance for creating novel eco-feedback systems based on real time consumption and grid power

generation data. This is particularly valuable because literature lacks designs that explore the disruptive changes emerging from the convergence of economic, technological and environmental factors. This study is further relevant because it provides actionable design guidance for creating novel eco-feedback systems based on mixed real-time production and consumption information.

WISE explores a large ecosystems taking advantage of a closed circuit energy system of an island in Europe with more than 270 thousand inhabitants. The average yearly electrical distribution is as follows: 78% of all energy is produced from thermoelectric plants, 11% from hydroelectric stations, 9% from wind parks and 2% from dispersed photovoltaic sources. We argue that our deployment in medium-sized isolated grid anticipates several issues that go beyond micro-generation scenarios and where the complex balance between the production / consumption reality and the individual / collective behaviours are closer to the future scenarios we might envision in the evolution of larger grids.

3 Interaction

Participants could interact with the Watt-I-See installation in three modes:

- *Physically engaging with the installation*: participants have to pedal on an exercise bike to surpass a micro-generation step to "power up" the installation. The energy produced does not influence the installation or the data it displays, rather its goal is to contextualize one's physical effort in producing electricity through a questionnaire on a tablet lying on the bike handle (See Figure 1 left).
- *Acknowledging the visual interface*: learning about percentages of renewable versus fossil energy being produced in that moment on the island. The WISE presents viewers with the real time values for the electricity produced locally.
- *Querying the interface of the installation*: the third mode of interaction allows participants to construct different days through cards representing several weather conditions and time-of-day. The visualization is then updated with the view its corresponding (real data) production quotas for those conditions.

3.1 The Vortices: Disaggregated Production Quotas

WISE resorts to an analogy of "x-raying" a household wall, displaying four glass pipes containing a colored vortex, each representing a energy production

source available locally: thermoelectric power stations, wind parks, hydroelectric stations and photovoltaic. The glass pipes contain distilled water and liquid paraffin. Each vortex is colored in order to represent a different energy source: dark purple for thermal energy, clear color for wind, dark blue for hydro and yellow for solar energy source. The size of the vortex, ranges from very low to very high in nine levels representing different quotas in percentage. The size of the vortices is measured from the top of the tube 0cm to the bottom 30cm. The highest level creates a more aggressive vortex to represent over 91% (limit selected by average maximum thermal quota) of quota from an individual production source. These levels represent the quota of energy produced and available to final consumers, thus the sum of all four vortices totals 100% (See Figure 1 right).

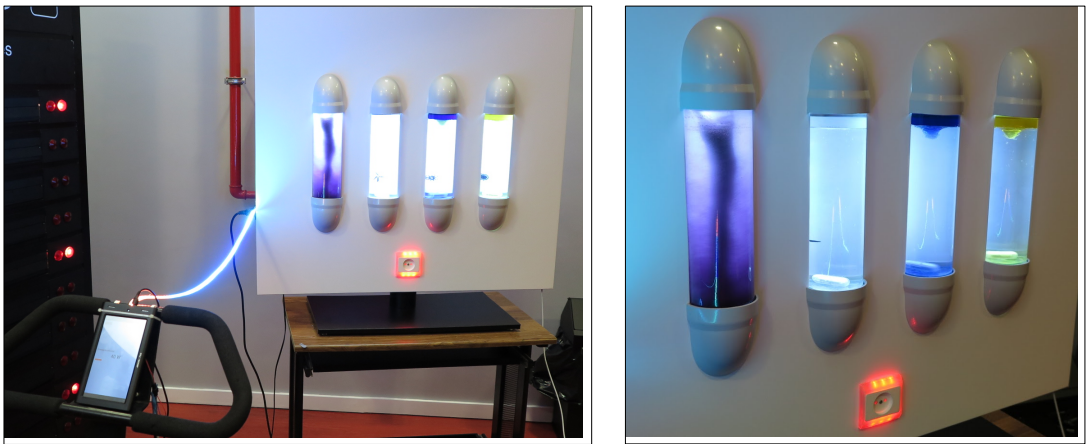


Figure 1: Left: Set up for the study with the exercise bike and the panel with the vortices. Right: Fossil source over 91%, 0% wind, hydro and solar sources between 1-3%. The socket is glowing red.

3.2 The Glowing Socket: Renewable Energy Feedback

In addition to the vortices, the power socket provides additional feedback on the overall quota of renewable energy in the grid. While the glass tubes and vortices display individual source production quotas, the power socket displays the cumulative quota of renewable energy in the power grid. Five renewable feedback levels were defined based on three years of disaggregated production quotas. Each level displays on the power socket a corresponding pulsing color (See Figure 2).

4 Implementation

4.1 Grid Electricity Production Data

The electricity production data, updated every fifteen minutes, is obtained directly from the regional energy provider currently the primary entity responsible for the distribution of electricity in the island. Additionally, the regional provider distributes, through a prediction model, the estimate production quotas for the next 12 hours.

4.2 Production Visualization

Watt-I-See measures 1.22m x 0.9m built from wood and covered in matte white vinyl. Four glass tubes (9x30cm) are used to represent the individual production sources. Vortices are created by DC motors controlled by an Arduino microcontroller. The DC motors rotate a large magnet that subsequently rotates a magnetic bar inside the glass tube creating the vortex. A smaller magnet and Hall effect switch is used to calculate the rpm's of each motor in order to leverage the rotation between the different tubes. Additionally, LED strips and drivers are used to retro-illuminate the glass tubes that are covered in tracing paper and contain the printed icons for each energy source.

The electricity socket is a common power socket where the outer bezel was 3D printed using a transparent PLA. The color of the socket is obtained by RGB LEDs and is based on several conditions queried in real-time to the database: 1) current real-time production quotas; 2) current day averages; 3) week averages; 4) month averages; 5) five hour prediction quotas. When providing feedback to consumers (or suggesting usage), WISE queries the database for the updated production data, to display through the vortices, and queries the database for a feedback value (between 1 red, and 5 green). An algorithm, which takes into account current and historical productions values, and the forecast for the rest of the day, returns a weighted value that measure how "green" the current production values are. The weighted value is finally compared to the overall three-year production averages, where minimums, maximums, averages and standard deviations were calculated to define the five feedback regions (see Figure 2).

An application was built in Processing in order to interface the Arduino and the web services allowing for additional interaction, such as choosing specific weather conditions and viewing real production data for those conditions and interaction with the exercise bicycle.

4.3 Micro generation

The micro generation step was performed through a repurposed exercise bike. A stepper motor was attached along with a rectifier circuit (converts AC to DC current) and voltage divider to limited voltage to 5V (input to an Arduino microprocessor) - which is responsible for calculating the electricity generated. An android tablet placed on the handlebars displays the progress. Additionally a digitally addressable (60) LED provides interaction feedback, such as: 1) standby status; 2) production feedback with a flow of energy effect and simultaneous progress bar; 3) goal attained; 4) regress/reset.

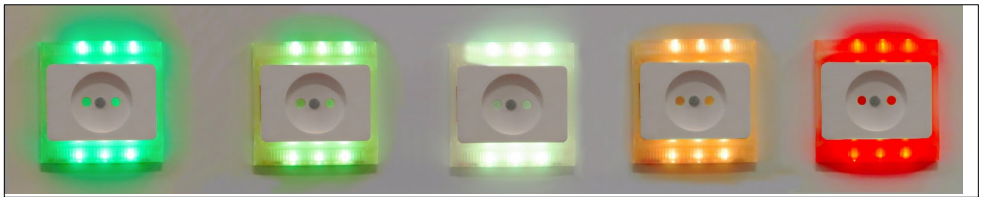


Figure 2: The power socket's five renewable energy feedback levels, from green (over 53% renewable) to red (less than 9%)

5 Results and Discussion

We conducted a preliminary evaluation of WISE using a combination of surveys, observations and interviews before and after 10 participants interacted with it. Figure 1 left displays the set up for this study. Next we present our major observations after all our interactions with participants where transcribed and organized according to the recurring themes.

5.1 Energy Literacy and Renewable Awareness

Being able to observe the values of the local energy production clearly increased participants' knowledge of the local grid and the efforts used to produce electricity. This observation has been called an increase in energy literacy [8] and is the result of an increase consumer's awareness.

5.1.1 Seeing is Believing

Results from the New Ecological Paradigm survey point out to consumers' awareness about the sources of energy, the different generation techniques and natural resources. However, during the study a majority of participants overestimate the impact of some renewable sources such as solar energy and widely underestimate the presence of hydroelectric and wind energy in the grid. Energy awareness, especially pertaining to renewable sources, seemed to

have an almost direct relation to participant exposure to the mechanisms used to produce them.

5.1.2 Memorable Events Incite Recall

The follow up survey, performed 8 to 10 days after the short exposure to WISE, found improved energy awareness results. Respondents were more aware of the source of their electricity and showed improved knowledge regarding how local electricity is produced and which sources are present in the grid.

5.1.3 Energy Production Efforts

The micro-generation procedure found a general lack of knowledge on the efforts needed for producing the electricity for domestic consumption. Low power longer usage appliances and devices seemed easier to understand and produced more correct answers, while high power appliances such as kettles yielded the most number of incorrect assumptions. It was clear that the notion of electrical power (watt hour) was complicated to grasp.

5.2 Exploring Materializing Energy through Movement

The simplistic representation of the energy production quotas thought the size of the colored vortex allowed for an attractive, understandable, almost mesmerizing effect, that was suggested to be faster to check and more intuitive. A stronger impact at times was noted. People felt somewhat worried when viewing the fossil vortex at its maximum scale how aggressive it looked. Others commented on the difference between fossil and renewables as depressive, something that needed to be dealt with.

5.3 The Importance of Feedback Immediacy

One of the preferred aspects of WISE was feedback immediacy. Providing immediate feedback was found as a two-fold design concern:

- Immediacy of feedback: Easy to understand, “at a glance” information. The glass tubes and vortices provide just enough information to inform a decision.
- Immediacy of interaction: feedback power socket is as close to the point-of-interaction as possible, however, possible at times not the most visible due to the location of power socket.

6 Conclusion

In this position paper we argue that HCI contributions in the domain of energy awareness and sustainable energy consumption, in particular, conventional ecofeedback, are detached from the wider political and economic contexts in which the utility business is evolving. The various disruptive challenges facing public utilities have different implications, but they all create adverse impacts on revenues, investor returns and ultimately energy price and usage and hence environmental impact. The WISE installation is a first attempt to physically represent grid production sources and quotas with an overall feedback mechanism. As such it was generally well accepted by our sample of users. Our preliminary study showed evidence about increased awareness and stimulated a dialogue about the different sources of energy, their relationship with weather and other context conditions and finally the consumption patterns in households. Here we attempted to provide some preliminary design insights for creating eco-feedback displays that strike a complex balance between the awareness of the realities and complexities of user lifestyles on one hand, and a desire to influence their use and consumption behaviours and practices on the other.

References

[1] Barreto, M., Karapanos, E., and Nunes, N. Social Translucence as a Theoretical Framework for Sustainable HCI. In P. Campos, N. Graham, J. Jorge, N. Nunes, P. Palanque and M. Winckler, eds., *Human-Computer Interaction – INTERACT 2011*. Springer Berlin Heidelberg, 2011, 195–203.

[2] Dourish, P. HCI and Environmental Sustainability: The Politics of Design and the Design of Politics. *Proceedings of the 8th ACM Conference on Designing Interactive Systems*, ACM (2010), 1–10.

[3] Filonik, D., Medland, R., Foth, M., and Rittenbruch, M. A Customisable Dashboard Display for Environmental Performance Visualisations. In S. Berkovsky and J. Freyne, eds., *Persuasive Technology*. Springer Berlin Heidelberg, 2013, 51–62.

[4] Froehlich, J., Findlater, L., and Landay, J. The design of eco-feedback technology. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2010), 1999–2008.

[5] Gustafsson, A. and Gyllenswärd, M. The power-aware cord: energy awareness through ambient information display. *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, ACM (2005), 1423–1426.

[6] Pierce, J. and Paulos, E. Materializing energy. *Proceedings of the 8th ACM Conference on Designing Interactive Systems*, ACM (2010), 113–122.

[7] Pierce, J. and Paulos, E. The local energy indicator: designing for wind and solar energy systems in the home. *Proceedings of the Designing Interactive Systems Conference*, ACM (2012), 631–634.

[8] Schwartz, T., Denef, S., Stevens, G., Ramirez, L., and Wulf, V. Cultivating Energy Literacy: Results from a Longitudinal Living Lab Study of a Home Energy Management System. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2013), 1193–1202.

[9] Strengers, Y.A.A. Designing eco-feedback systems for everyday life. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2011), 2135–2144.

[10] *International Energy Outlook 2013*. United States Energy Information Administration, 2013.

Illustrating Energy Related Properties of Buildings Using a 3D-Game-Engine

Thomas Rist, Jens Müller
University of Applied Sciences Augsburg
Augsburg, Germany
{Thomas.Rist, Jens.Mueller}@hs-augsburg.de

Abstract. The paper presents first steps towards the deployment of a 3D-game engine for the interactive exploration of energy-related properties of buildings, such as heat transfer properties of different materials, or the effect of settings of HAVAC (Heating, Ventilation, and Air Conditioning) control systems on a building's in-door climate. We present a number of relevant information goals in the domain of energy efficient buildings and relate them to illustration techniques which could be used to increase a user's understanding of and interest in the topic. It is assumed that a repertoire of visually attractive illustration techniques will pave the way for novel game-like exploration systems for non-expert users in the domain of energy-efficient architecture and construction.

1 Motivation

Energy efficient buildings are becoming increasingly important given the desired political and societal shift away from fossil fuels. Various tools have been developed to assist practitioners and researchers in the domains of architecture and construction engineering to explore and evaluate energy-efficiency technologies and renewable energy strategies in new or existing buildings. For an impressive list of more than 400 such tools see the catalogue maintained by the Office of Energy Efficiency and Renewable Energy of the US Department of Energy [1]. However, designed for professional users, such tools are often too complicated and too expensive for a broader user group including potential private home builders as well as owners who want to inform themselves about existing and upcoming energy-efficiency technologies. The importance of this target group is given by its sheer size, though variations exist from one country to another [2].

Our aim is to allow non-expert users to access and experience in a playful way energy-related properties of buildings, construction materials, and building usage patterns by means of a popular 3D game engine (e.g. Unreal or Unity).

For the purpose of a case-study we cooperate with the principle constructor (i.e., a consortium consisting of the council of the German town Königsbrunn, a regional energy services provider, a company for residential development, and our university) of a highly energy efficient prototype building, the "Visioneum" [3]. The Visinoeum is a so-called energy plus building, i.e., a building that over the year produces more electric energy through roof-top solar power panels than it consumes for heating/cooling, lighting, and electric home appliances. The surplus of electric energy may be either fed into the power grid, or likewise used for fuelling electric vehicles in its neighbourhood. When completed in 2016, the Visinoeum will be open to the public and host an exhibition space as well as an office for energy consultancy. In addition, our university will use it as a test-environment for research on new energy efficient materials and building automation technologies. For our project on interactive 3D illustration of energy- related properties we take advantage of the Visioneum's 3D CAD data.

2 Approach

In our attempt to explain and illustrate energy relevant properties of a building, its materials, and usage patterns we need to bring together expertise from different disciplines, building design and engineering, 3D illustration, and computer game development.

2.1 Identifying Information Goals with Domain Experts

In a working meeting with colleagues from the department of construction engineering we compiled a list of themes which are of special interest when discussing energy performance of buildings; among those themes are:

- Heat transfer and insulation properties of building parts (such as walls, ceilings, staircases, windows, doors etc.) depending on the materials of which they are made of, as well as of their structural properties, such as surface area.
- Functioning and effectiveness of the HAVAC (Heating, Ventilation, and Air Conditioning) control system of an inhabited building at different seasonal requirements.
- Orchestration of solar energy generation, battery-based storage, and smart consumption by means of an intelligent building control system.
- The integration of an energy plus building in a neighbourhood which consists of conventional as well as further energy plus buildings.

- Sustainable, resources-preserving building design, use of renewable materials, and recyclability after demolition.

Next, the domain experts formulated questions and information objectives related to afore mentioned themes. Thereby, identified information goals go beyond mere explanations of how certain technologies work.

Regarding properties of building parts raises the question of what are the effects when a part gets replaced by another part made of other materials. For example, different materials for walls and windows affect heat transfer but there may be additional effects as well, such as a change in noise dampening or the need for modifications in the building's overall construction. Thus, making a user aware of consequences and trade-offs caused by material choices is an important information goal in this context. In the case of the air ventilation system the information goal could be to make the user aware about possible consequences of human interventions, such as opening a window.

2.2 Development of Interactive Illustration Techniques

Technical illustration, a discipline at the cross section of visual art, science, and engineering, has brought about a multitude of visualisation styles, rendering techniques, and visual metaphors to visually communicate technical subject matters to a nontechnical audience. Inspirations and guidelines for the production of handcrafted illustrations can be found in textbooks [4], [5], [6], as well as in online collections [7], [8]. Popular types of technical illustrations are:

- annotations of objects, e.g. by text labels to name depicted object;
- exploded views, e.g., to show the spatial arrangement of the parts of an assembly;
- cut-away views and transparent views, e.g., to enable a look through or inside non-transparent objects;
- cross-sections, e.g., to show the structure of layered materials;
- ghost or phantom objects, e.g., to show selected locations of an object movement,
- variations of level of detail, e.g. to draw the viewer's attention from background objects to foreground objects;
- non-realistic and false-colour rendering of surfaces, e.g., to encode otherwise invisible object properties, such as colour coded heat-maps an object's surface temperature distribution, or to emphasise or deemphasize certain object parts through artificial shading effects, such as hatching;

- inclusion of meta-graphical objects, e.g., arrow symbols to indicate the direction of a moving physical object, or lines to visualise sound waves or energy fields.

The above mentioned techniques have been developed for the purpose of static printable technical documentation. Nowadays, the production of technical illustrations benefits from powerful graphics software packages. Nevertheless, the majority of technical illustrations are still static images even if distributed digitally and viewed on computer screens, or they are rendered as a sequence of images to be viewed as a video short.

In contrast, interactive illustrations are dynamic media as they allow viewers to explore depicted subject matters through instant changes of view points, and modifications of object properties. Interactivity is brought in by means of a user interface that enables variation of parameters to control type, style, and display of an illustration. In some approaches parameters have to be set prior to the display of the illustration. For example, a user may set parameters for an automated camera-flight around or into an object (or building), such as start and end point or the speed of the camera movement. Other settings allow for more direct control, such as user-controlled 3D walk-throughs, or instant interventions, such as manipulating objects. Within the discipline of computer graphics interactive versions for all of the above mentioned techniques have been developed. Examples include the dynamic rearrangement of textual annotations [9], modification of render style [10], or cut-away views [11], [12] while performing view point changes.

So far, only a few attempts have been made to exploit techniques of interactive illustration to the domain of energy-efficient buildings. Schreyer and Hoque [13] use thermography images of buildings taken by an infrared camera as textures for rendering a simplified 3D model of the corresponding building. This way, a series of thermography images can be explored as a virtual 3D tour around the building. Work at Autodesk Research [14], [15] combines information of a BIM (Building Information System) with measured data collected through a sensor network to better understand a building's performance during usage. They have developed a number of 3D-techniques for illustrating heat distribution and heat flow within an office space. For example, gradient shading is applied to surfaces to visualize thermal values of selected spatial zones. An interesting approach to visualize sensor recordings has been proposed by Wittenburg et al. [16]. They represent time series of sensor readings in the shape of sliced 3D-scupltures which are positioned on a 2D floor plan to show the locations of the sensors in a building.

Our approach shares similarities with work conducted in the EEPOS project [17]. As part of the user interface for neighbourhood energy service providers

and users they integrated the Unity game engine. Users can explore 3D models of apartments and receive real-time as well as historical data of energy consumptions through text and chart displays which are integrated into the 3D model. However, our approach is not on building monitoring and thus goes beyond the display of energy consumption or generation. Rather, users are invited to modify building properties and learn about the effects by means of interactive illustration.

2.3 Adding Elements of Game Play

To familiarise non-expert users with topics related to energy-efficiency of buildings we believe that adding elements known form computer games can be beneficial, since one can draw on:

Users' likely familiarity with computer games: Many people have already played computer games and are thus already familiar with popular principles of game play and interaction styles found across different computer game genres. Our Visioneum show case features a 3D environment to be explored by users through navigation in virtual 3D space with a free-roaming camera, manipulation of objects in the environment, and triggering events that produce perceivable outcomes. Such elements are found in a number of game genres, including real-time action, adventure, and maze games. Also, we expect that users will easily understand how to control the application using a gamepad.

Users' intrinsic motivation to play: For leisure games Ryan et al. propose "… that events and conditions that enhance a person's sense of autonomy and competence support intrinsic motivation, whereas factors that diminish perceived autonomy or competence undermine intrinsic motivation." [18, p 349]. Many serious games approaches draw on the assumption that embedding a learning task -such as teaching a technical subject-matter- into a game context can motivate players to engage in them similar as leisure games do. However, a recent meta-study by Wouters and colleagues did not confirm an increased motivational appeal of serious games compared to other learning methods per se [19]. As a possible reason they point out that learning games are often part of a curriculum and therefore give players less autonomy with respect to when they play and for how long they want to play. Fortunately, this factor will be less relevant in our Visinoneum showcase as the visualisation system is for use on a voluntarily basis.

3 State of Development and Outlook

Our project is still at an early stage of development. So far, we extracted from the Visioneum's CAD data a simplified model suitable for import into the Unreal 4 graphics engine (cf. Fig. 1 a). Starting from the information goals identified by the domain experts, we are currently experimenting with different types of interactive visualisation techniques. For example, to explain the effect of replacing a window with another one that has other insulation properties, we may deploy particles to visualize heat flows.

Our current implementation does not yet constitute a (serious) game. However, we intend to use it as a test-bed for adding elements of game play. For instance, the goal to inform a user about the effect of different materials insulation could be transformed into a player's "mission" to detain as much as possible heat particles in the building. To do so, the player would have to pick appropriate materials for the building parts, and while doing so, learn about their properties and possible constraints on their combinability.

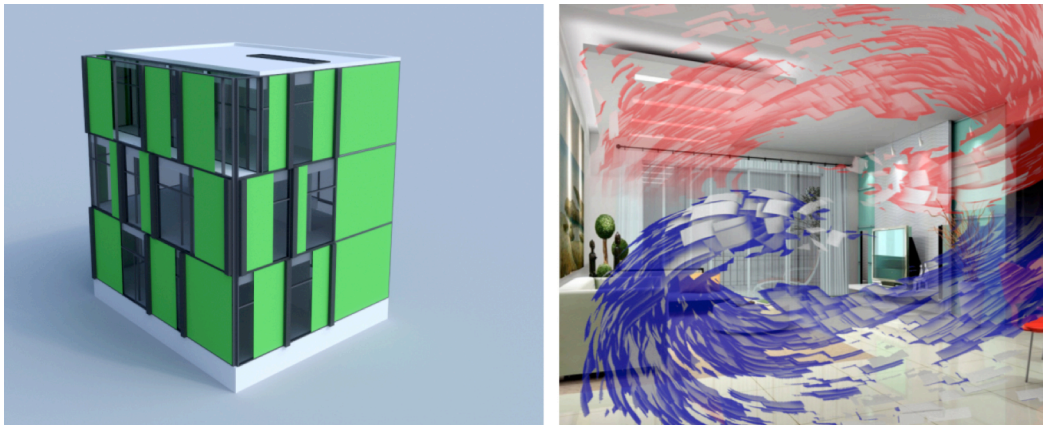


Fig 1: (a) 3D model of the Visioneum, (b) Particles forming 3D- arrows to depict air circulation in a room

Further inspirations for adding game elements can be found in energy games [20]. A recent study by Grossberg and colleagues surveyed 53 games which aim to influence behavior around energy efficiency and sustainability [21]. Focusing on 22 games that are or could be part of an energy efficiency program, they discuss how these games address essential game elements, such as progress paths, levels, and triggers, real-time feedback, interaction and competition among players, and achievements and rewards. To meet the preferences of different player types [22], there may be a need to offer a choice between different exploration styles.

In addition, we plan to evaluate our hypothesises that users will (a) quickly be able to familiarize with the exploration system, and (b) appreciate the exploration experience.

Acknowledgements

We like to thank the members of our project group Kevin Aust, Dennis Falkner, Markus Heimbach, Ines Kossack Marlen Materna, Alexander Mersdorf, Carmen Merz, and Martin Willam. Special thanks goes to our colleagues from the department of architecture and construction engineering Georg Sahner and Johannes Rieger for their valuable inputs. The work is supported by the BMBF-funded German-New Zealand research collaboration IT4SE, grant number 01DR12041. Further information on IT4SE can be found under www.it4se.net

References

[1] Office of Energy Efficiency and Renewable Energy (EERE). Building Energy Software Tools Directory: http://apps1.eere.energy.gov/buildings/tools_directory/
[2] Wikipedia.org. List of countries by home ownership rate. http://en.wikipedia.org/wiki/List_of_countries_by_home_ownership_rate
[3] Visinoeum Energie+: http://www.visioneum.de
[4] Thomas, T.A. Technical Illustration. McGraw-Hill, 1960.
[5] Martin, J. Technical Illustration: Materials, Methods and Techniques , MacDonald Orbis, 1989.
[6] Yee, R. Architectural Drawing: A Visual Compendium of Types and Methods. John Wiley & Sons; 4th edition, 2012.
[7] Technical Illustrators: http://technicalillustrators.org/
[8] Pinterest Inc.: http://www.pinterest.com
[9] Ali, K., Hartmann, K., and Strothotte, T. Label Layout for Interactive 3D Illustrations. Journal of the WSCG, 13:1–8, 2005.
[10] Gooch, B., Sloan, PPJ., Gooch, A., Shirley, P., Riesenfeld, R. Interactive technical illustration. Proceedings of the 1999 symposium on Interactive 3D graphics, 1999.
[11] Feiner, S., and Seligmann, D. Cutaways and Ghosting: Satisfying Visibility Constraints in Dynamic 3D Illustrations. The Visual Computer, 8(56), 1992, pp. 292-302.
[12] Diepstraten, D., Weiskopf, and Ertl, T. Interactive Cutaway Illustrations. Computer Graphics Forum, 22(3):523–532, September 2003
[13] Schreyer, A.C., and Hoque, S. Interactive Three-Dimensional Visualization of Building Envelope Systems Using Infrared Thermography and SketchUp. Proceedings of InfraMation, Vol.9, 2009.

[14] Hailemariam, E., Glueck, M., Attar, R., Tessier, A., McCrae, J. Khan, A. Toward a unified representation system of performance-related data eSim 2010 Conference Proceedings: IBPSA-Canada eSim Conference. 2010, pp. 117 - 124.
[15] Attar, R., Hailemariam, E., Breslav, S., Khan, A., Kurtenbach, G. Sensor-enabled Cubicles for Occupant-centric Capture of Building Performance Data. Proc. of ASHRAE Annual Conf. 2011
[16] Wittenburg. K., Laughman. C., Nikoiski. D., Sahinoglu. Z. Advanced Visual Interfaces for Smart Energy: Focusing Where it Matters Most. Proc. of the AVI 2014 Workshop on Fostering Smart Energy Applications through Advanced Visual Interfaces. 2014, pp. 31-34
[17] EEPOS project website, http://eepos-project.eu
[18] Ryan, R. M., Rigby, C. S., and Przybylski, A. The Motivational Pull of Video Games: A Self-Determination Theory Approach. Motivation and Emotion, 30, 2006, pp. 347–365.
[19] Wouters, P., van Nimwegen, C., van Oostendorp, H., van der Spek, ED. A meta-analysis of the cognitive and motivational effects of serious games. Journal of Educational Psychology, Vol 105(2), May 2013, 249-265.
[20] Canada Science and Technology Museums Corporation. Energy Games. http://energy.techno-science.ca/en/games-videos/energy-games.php
[21] Grossberg, F., Wolfson, M., Mazur-Stommen S., Farley K., and Nadel, S. Gamified Energy Efficiency Programs American Council for an Energy Efficient Economy (ACEEE), 2015. Online: http://aceee.org/research-report/b1501
[22] Bateman, C., Lowenhaupt, R., Nacke, LE. Player Typology in Theory and Practice. Proc. of DiGRA 2011.

Simulating the Impact of Household Energy Consumption on the Electricity Grid

Patrick Ozoh, Mark Apperley
Department of Computer Science
University of Waikato
New Zealand
{pozoh,m.apperley}@waikato.ac.nz

Abstract. This paper presents a bottom-up modelling approach for stochastic production of electricity consumption profiles in households. It represents a preliminary work on individual appliance use modelling in households, as part of a bottom-up simulation to assess the impact of household consumption, and changes to consumption patterns and behavior, on the overall energy grid. By collecting household electricity consumption data, a model is developed based on daily activity profiles for individual appliances. The domestic load model obtained from simulating electricity consumption for household appliances will enable the large-scale simulation of multiple households to gain insight into individual household implications of demand-side load management strategies, as well as the combined effects on the electricity grid.

Keywords: Domestic appliance modelling, bottom-up model, stochastic, simulation

1 Introduction

Characterizing electricity demand profiles for households is an important prerequisite for analysing demand side management, (Richardson & Thomson, 2008). These researchers suggest that electricity demand is influenced by active occupancy, (see Figure 1). This graph shows very low levels of active occupancy during the night, delayed arising at weekends, and peaks corresponding to mealtimes - breakfast, lunch and dinner.

Because the use of household electric appliances varies with time, since understanding time-of-use profiles for producing household appliance use models based on various parameters such as location, socio-economic class, age groups and household size is necessary as an input to any domestic energy model that is to be used for assessing the effectiveness of demand-side load management on efficient energy, and its impact on comfort and quality of life at individual household, community, regional and national levels.

In this vein, the development of a model for generating residential electricity and hot-water load profiles from time-use data was developed by

Widén et al. (2009). Daily time-use data sets and energy measurements for different appliances were used to develop a model that can be used for the determination of households' energy consumption. The study collected data for each occupant over ten years old in participating households, recording at 5-min intervals, including a description of the activities, geographic location of the person, their means of transport (if relevant), and by whom they were accompanied while performing the activity. One application of the developed model is visualization of energy use associated with everyday activities. The study demonstrated that energy use profiles can be generated from time-use data measurements with reasonable accuracy when compared with actual electricity consumption measurements.

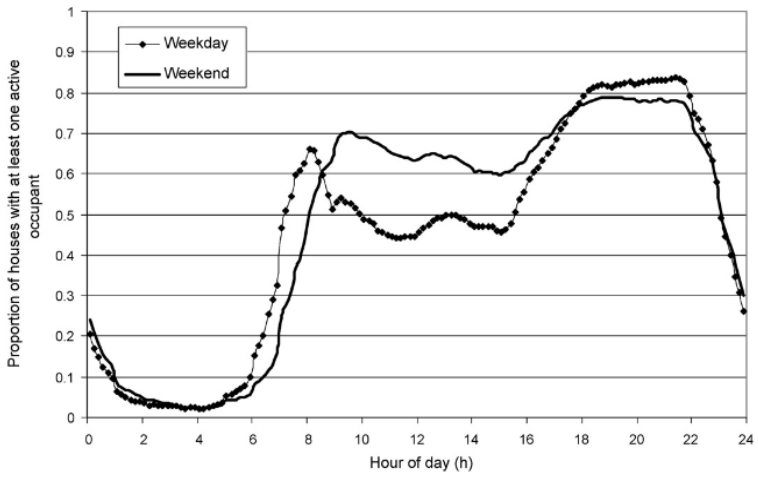
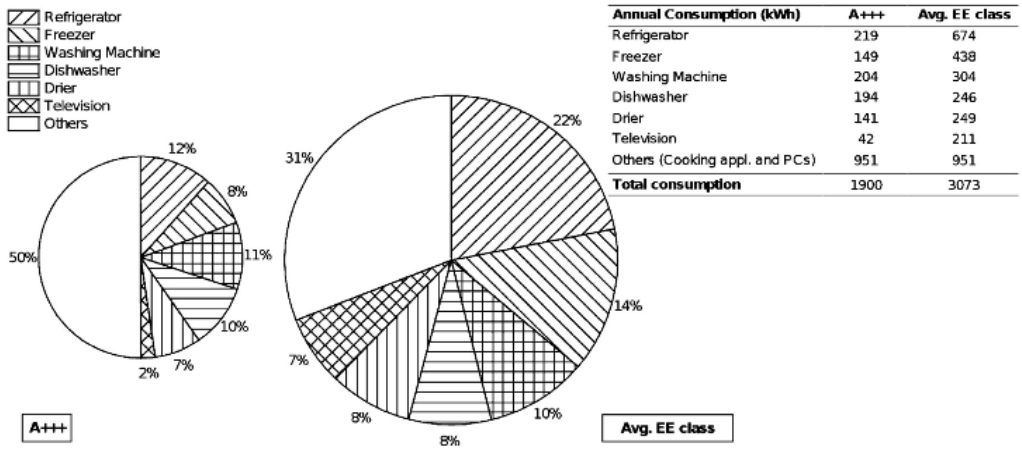


Figure 1: Aggregated active occupancy for households by weekday and weekends, (reproduced from Richardson & Thomas, 2008)

Grandjean et al. (2011) described the simulation of power consumption and the development of a domestic load curve model for individual electric appliances, in order to investigate their individual power demand. The study was motivated by the need for behavioural and lifestyle modifications of household's electricity usage as their energy consumption are likely to increase in the near future through new electrical domestic end-uses, including the plug-in hybrid and electric vehicles, increased heat pumps utilization, and improvement and technological advances in small electrical appliances. The research considered real power consumption measurements for individual appliances, and simulated the behaviour of households to develop their load profiles. This approach is referred to as a bottom-up modelling process.

A more recent description of household power consumption simulation (Ortiz, Guarino, Salom, Corchero, & Cellura, 2014) developed an energy consumption model for generating random profiles, useful for simulating

energy consumption for a cluster of buildings or individual appliances, as well as to model their peak loads. The stochastic model developed for this study reproduced random synthetic profiles for household consumption, applied in high time resolution aimed at modelling peak loads as well as producing average data for households. The research quantified the benefits of improving the energy efficiency of appliances by showing a reduction in energy consumption by half when high performance appliances were introduced (Figure 2). However, the study is highly reliant on the input data of energy consumption for the households under consideration. Dependency on data at this level places demands on data storage system and on continuous internet connectivity of households.



Annual Consumption (kWh)	A+++	Avg. EE class
Refrigerator	219	674
Freezer	149	438
Washing Machine	204	304
Dishwasher	194	246
Drier	141	249
Television	42	211
Others (Cooking appl. and PCs)	951	951
Total consumption	1900	3073

Figure 2: Comparison of household's appliance use and energy efficiency ratings of appliances (reproduced from Ortiz et al., 2014)

Another study of the consumption of electric appliances in domestic buildings was carried out to identify the trends in their energy use pattern, (Firth, Lomas, Wright, & Wall, 2008). In this research work, the consumption of different energy user groups (low, medium and high) was investigated. The study identified low and high power users as major contributors to total electricity consumed by appliances. Yao & Steemers (2005) discussed a simple method of formulating load profile (SMLP) for domestic buildings in the United Kingdom, in which electric appliance load profile and domestic hot water profiles were calculated by gathering input data for the daily average end-use energy consumption and daily average hot water consumption of households respectively. The study confirmed that load trend for households are close to the national statistics data. It is suggested that the SMLP method can help electricity suppliers predict the likely future development for household electricity demand.

A number of other reported studies discuss industrial and household energy consumption. An exploratory analysis of domestic electricity profiles recorded at a high time resolution, taken at one minute time intervals suggested usage pattern of households varies widely, with some larger loads requiring much greater energy supply, suggesting the need for the development of effective storage technologies (Wright & Firth, 2007). The potential for information feedback to reduce rates in the home through energy consumption indicators (ECI) or smart meters is explained in Figure 3 (Wood & Newborough, 2003). The study focused on collecting data from individual appliances and then compared the effectiveness of providing paper-based energy use/saving information with electronic feedback of energy-consumption via specifically designed energy-consumption indicators.

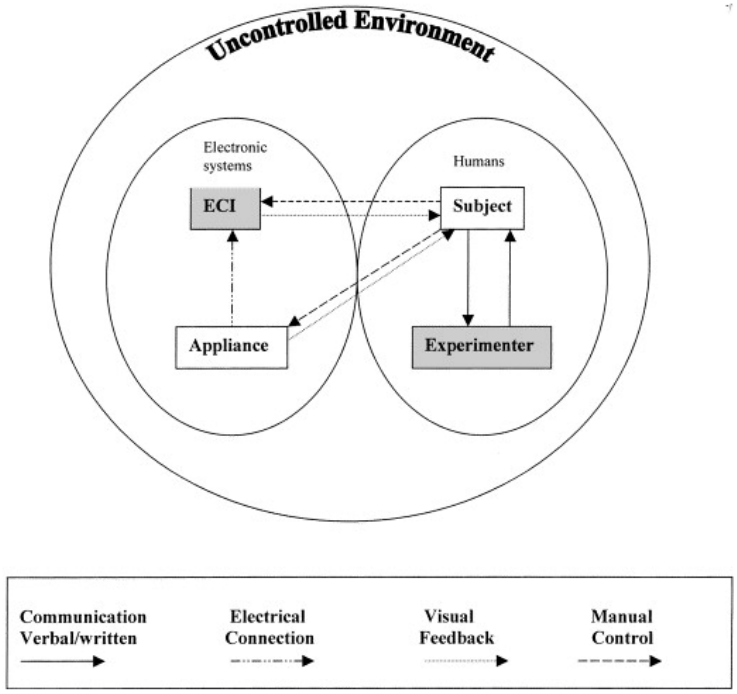


Figure 3: The feedback process for domestic appliance use (reproduced from Wood & Newborough, 2003)

Mansouri, Newborough, & Probert (1996) described the need to provide end-users with accurate energy-consumption and environmental impact information to stimulate energy saving and environmentally sustainable behaviour. The paper focused on identifying energy-use behaviour, environmental activities and benefits, ownership levels for certain appliances and their utilization patterns among households.

Various studies have discussed the development of load profiles for electric appliance use in households. However, most come with challenges and short-comings, such as complexity of the modelling and a lack of explicit

description of consumption at appliance level, and detailed time-of-use information which is essential for effective modelling of the grid impact of demand-side management strategies. In this paper, a stochastic model based on appliance-specific measurements in households is developed. The aim of this research is to keep the required input data at minimum and the model structure as simple as possible, to enable the measurement of appliance use impact on the electricity grid. In Section 2, data-sets of domestic load for appliances used in this study are described. Conclusions are drawn in Section 3.

2 Model Framework

Data-sets of daily load in kilowatts were collected from households in Hamilton, New Zealand between August 2011 and March 2012. Data was collected using monitors connected directly to individual appliances. Data quality for the measurements was generally high, although there were some gaps in the data, but only complete data-sets were used in the analysis. The time series analytical technique was used to model data collected. This method was selected because it models consumption data for individual appliances over time and investigates the behaviour of appliance-use data.

The pattern of appliance use for domestic loads in buildings depends on the appliance use. Some appliances are ‘always on’, while others appliances are not constantly on and their state depends on occupancy, occupant behaviour and weather conditions, which vary between households. Typically, electricity consumption by appliances is higher than usual in winter because of heating, and occupancy issues such as time spent indoors.

Figures 4(a) to (d) show load profiles for refrigerator, LCD-TV, microwave and play-station and follows an undulating pattern. The figures show the load-profile of electricity consumed by different appliances at different times of the day. The figure shows sharp increases in consumption for refrigerators at certain periods of the day. This could be due to defrosting of fridges in households, as this process consumes much power. There are times when power for appliance-use was low, because they were in ‘standby’ mode, i.e. appliances are not in use, but ready to go into operation.

The load profiles developed for individual appliance give information about habits of household inhabitants in their appliance use. Load shifting by consumers enables reduction of peak demand on the energy grid and increases the reliability of the electricity network, since consumers will have a more active role concerning electricity management.

An outline of electricity consumed by individual appliances is shown in Figure 5(a). The figure indicates a high consumption by the refrigerator (10%) compared to other appliances; LCD-TV was 2.7%, play station 2.1%, and microwave 1.2%. The remaining power consumption was for several other appliances in use consisting of 84% of total consumption. There were some fluctuations due to appliance use resulting in peak and off-peak periods, depending on whether appliances are in use. There is more near-constant load for LCD-TV and the play station except when the appliances were in standby mode.

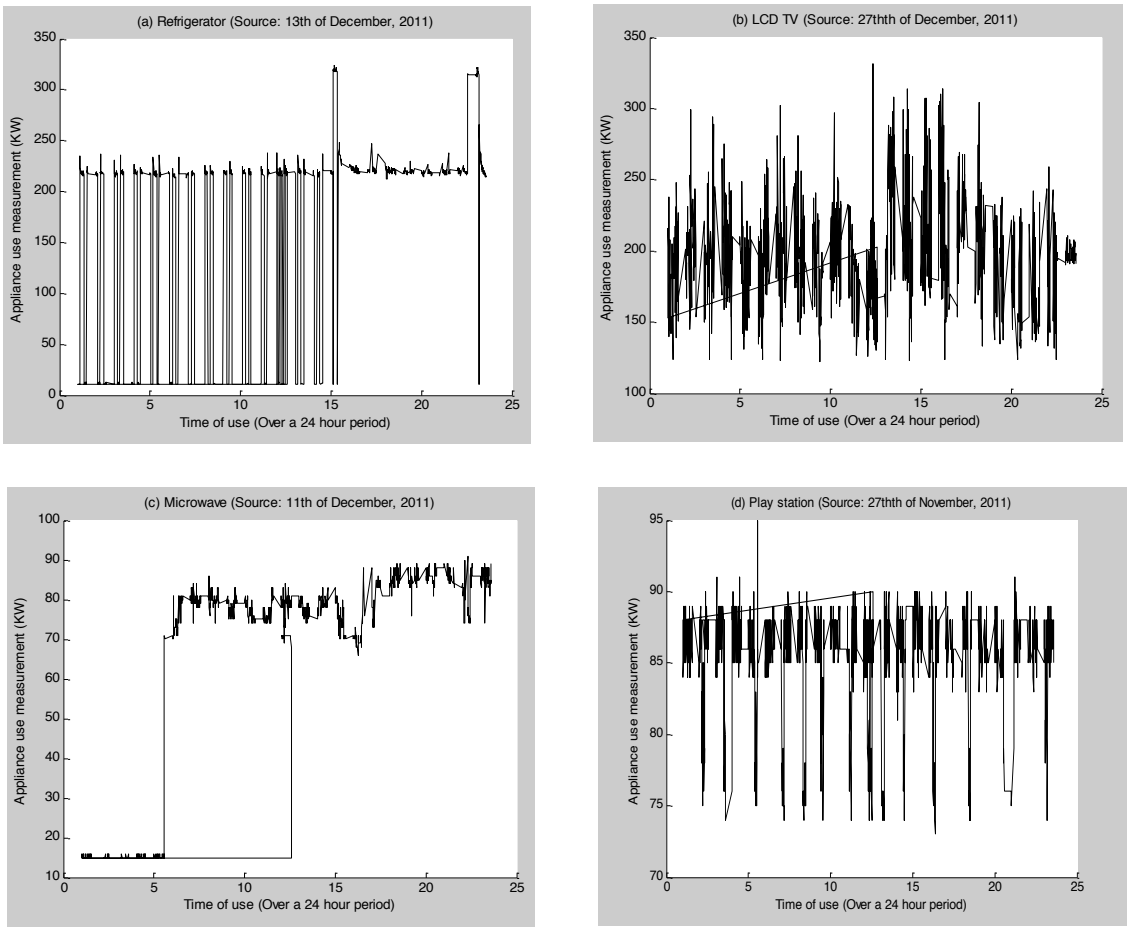
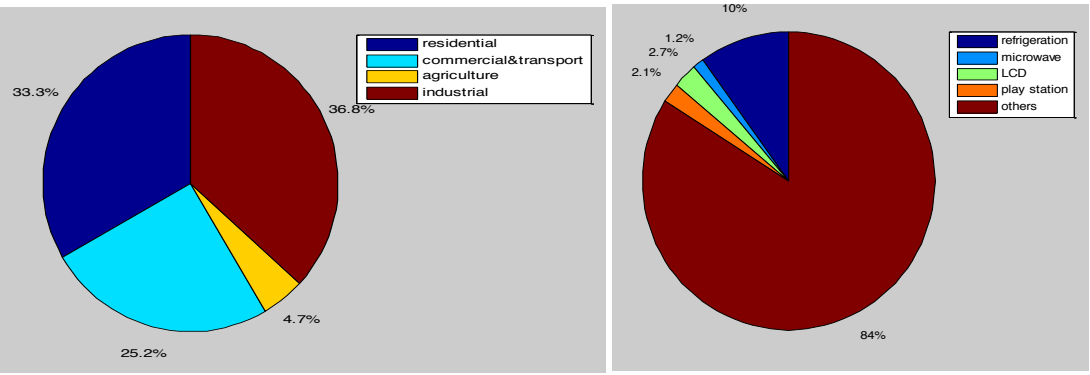


Figure 4: Daily one-minute interval load profiles for appliances (in kW)

There are several other electric appliances contributing to total consumption in individual households. Some of them include electric kettles, television sets, washing machines, dryers, iron, toasters, iPhone chargers, laptops, freezers, etc. These appliances constitute 'others' in Figure 5(a). A collection of appliances comprising of refrigerators, microwaves, LCD, and play stations were identified and used in this study because these appliances were in use in most households used for data collection.

The load profiles of the refrigerator, microwave, play- station and the LCD-TV were developed in order to investigate the impact on the energy grid of consumer use of electric appliance during peak and off-peak hours. Electricity consumption in households is about 33% of the national grid load in New Zealand (Ministry of Economic Development, 2012); see the relationship between Figures 5(a) and 5(b).



(a) Household appliance consumption (Source: DUMP project)

(b) Energy use in the electricity grid (Source: NZ energy data, 2013)

Figure 5: Relationship between residential appliance use and sectors in the electricity grid

3 Conclusion and Future Work

This paper describes work in progress to model household individual appliance use in order to assess the impact of different demand-side strategies on individual household consumption, and the energy grid. The development of load profiles for individual household appliances enables investigating the impact of load shifting strategies aimed at reducing energy costs and the optimal management and control of the electricity grid. Furthermore, the introduction of time-of-use energy plans by energy companies to better manage electricity consumption will enable households reduce their electricity bill, become responsive to using their appliances in a more efficient manner, and economically conscious users, thereby making optimal use of the energy grid. This will bring about load balancing and optimal use of the electricity network by redistributing energy demand in the electricity network.

Investigative research carried out in this study suggests that data on socio-economic factors is relatively sparse and conducting more research on the effect of socio-economic factors on household electricity consumption would help in providing insight to household appliance use. The next stage of research involves collecting data from households from different location/regions and will focus on developing load profiles for different group

of users based on location, income level and household size. The research will also investigate the effects of load shifting strategies on individual appliances based on these factors.

References

[1] Firth, S., Lomas, K., Wright, a., & Wall, R. (2008). Identifying trends in the use of domestic appliances from household electricity consumption measurements. Energy and Buildings, 40, 926–936. doi:10.1016/j.enbuild.2007.07.005

[2] Grandjean, A., Binet, G., Bieret, J., Adnot, J. & Duplessis, B. (2011). A functional analysis of electrical load curve modelling for some households specific electricity end-uses, 6th International Conference on Energy Efficiency in Domestic Appliances and Lighting (EEDAL '11) Copenhagen, 24p.

[3] I. Richardson & M. Thomson, D. I. (2008). A high-resolution domestic building occupancy model for energy demand simulations. Journal, Energy and Buildings, 40(8), 1560–1566.

[4] Mansouri, I., Newborough, M., & Probert, D. (1996). Energy consumption in uk households: Impact of domestic electrical appliances. Applied Energy, 54(3), 211–285. doi:10.1016/0306-2619(96)00001-3

[5] Ministry of Economic Development. (2012). New Zealand Energy Data File 2012, 172.

[6] Ortiz, J., Guarino, F., Salom, J., Corchero, C., & Cellura, M. (2014). Stochastic model for electrical loads in Mediterranean residential buildings: Validation and applications. Energy and Buildings, 80, 23–36. doi:10.1016/j.enbuild.2014.04.053

[7] Widén, J., Lundh, M., Vassileva, I., Dahlquist, E., Ellegård, K., & Wäckelgård, E. (2009). Constructing load profiles for household electricity and hot water from time-use data—Modelling approach and validation. Energy and Buildings, 41(7), 753–768. doi:10.1016/j.enbuild.2009.02.013

[8] Wood, G., & Newborough, M. (2003). Dynamic energy-consumption indicators for domestic appliances: Environment, behaviour and design. Energy and Buildings, 35, 821–841. doi:10.1016/S0378-7788(02)00241-4

[9] Wright, A., & Firth, S. (2007). The nature of domestic electricity-loads and effects of time averaging on statistics and on-site generation calculations. Applied Energy, 84, 389–403. doi:10.1016/j.apenergy.2006.09.008

[10] Yao, R., & Steemers, K. (2005). A method of formulating energy load profile for domestic buildings in the UK. Energy and Buildings, 37, 663–671. doi:10.1016/j.enbuild.2004.09.007

Interactive Solar Panel Simulation Tool - From GHI to PV Output

Joris Suppers, Mark Apperley
Department of Computer Science
University of Waikato
Hamilton, New Zealand
jorissuppers@hotmail.com

Abstract. The production and integration of solar panels is steadily growing in New Zealand and worldwide. The most common way to install solar panels in New Zealand is flat on top of the most north-facing roof. This method is used as it is space efficient, reduces mounting costs and will capture most of the sun's movement throughout the day. New Zealand homes typically have a small roof pitch. This means an installed solar panel will have a low tilt, allowing for best input in summer. However, this may not be the best result for individuals, as they may want a different outcome such as: maximum energy output in winter, maximum yearly output or a balanced output throughout each month. This paper describes a solar panel tool, which has been developed to enable users to explore interactively the impact on electricity production based on the orientation of the panels.

1 Introduction

This paper explains a solar panel simulation tool developed (nzspot.cms.waikato.ac.nz) that is designed to be interactive and enable individuals to easily see the effects on power output when changing the solar panel(s) orientation in real time. This allows the user to determine the best configuration to meet their criteria. The three stages of this simulation tool are: (1) selecting the location of the solar panel(s), (2) choosing the solar panel setup and inverter, and (3) setting the tilt and orientation of the solar panel(s). The paper then concludes by comparing the effects of different orientations on power output.

2 Location

The first stage of this tool is to determine the location of the solar panels. The user is presented with a map (Google maps) focused on New Zealand, allowing the individual to easily navigate and select their location. Once the location is picked, the tool retrieves the latitude and longitude values. With these values it is

possible to estimate the sun's position, irradiance, wind speed and dry bulb temperature for anytime during the year.

2.1 Sun's position

Knowing the location of the sun is essential in estimating a solar panel's power output as it directly relates to the amount of irradiance that will hit the solar panel at a given time. The sun's position from the earth's perspective changes throughout the day because of the earth's axis of rotation and the path it follows around the sun. This tool calculates seven different variables to accurately estimate the sun's position at any time. These variables include: declination angle, equation of time, solar time, hour angle, elevation angle, zenith angle and azimuth angle.

2.2 Irradiance

With the sun's position and the solar panels location known it is possible to estimate the irradiance that will reach this location. This irradiance includes global horizontal irradiance, direct normal irradiance and diffuse horizontal irradiance.

2.2.1 Global Horizontal Irradiance (GHI)

Global Horizontal Irradiance is the total amount of radiation hitting a surface horizontal to the earth's surface. This includes both direct normal irradiance and diffuse horizontal irradiance. The tool uses the location information provided to retrieve measured GHI from the nearest weather station. The data for these weather stations is retrieved from The National Climate Database from NIWA (cliflo.niwa.co.nz).

2.2.2 Direct Normal Irradiance (DNI)

Direct Normal Irradiance describes the amount of solar radiation hitting a surface perpendicular to the sun's rays. This simulation tool uses a model developed by Laue (Laue, 1970) to estimate DNI. This model is based on air mass and a clearness index of 0.7. The 0.7 is an approximation of clearness index, since roughly 70% of solar radiation is transmitted to the earth's surface. This tool calculates the real clearness index so it is used instead.

2.2.3 Diffuse Horizontal Irradiance (DHI)

The diffuse horizontal irradiance describes the amount of solar radiation hitting a surface, which is not from direct sunlight e.g. scattered and diffused radiation. Using the equation by Maxwell (Maxwell, 1987) DHI can be estimated.

2.2.4 Wind speed and Dry Bulb Temperature

Along with GHI a further two variables that this tool retrieves from the nearest weather station are: wind speed and dry bulb temperature. These values are needed in future calculations to get a good estimate for cell and module temperature.

2.3 Summary

The first stage of this tool is simply retrieving the location of the solar panels. This tool achieves this by presenting the user with a map and allowing them to click on a location. With this location known the tool then retrieves the global horizontal irradiance, wind speed and dry bulb temperature from the closest weather station, and calculates the irradiance hitting a surface at this location. This is all done in the background to minimize the user input to achieve a simplistic and easy to use website.

3 Inverter and Module Configuration

The second step of the simulation tool is to specify the characteristics of the module (solar panel) and inverter. This tool allows users to enter their own module and inverter characteristics, or for simplicity search over a large database from Sandia labs[1], which contains hundreds of inverters and modules. This searching is simply done by clicking the specific combo box; this will present a variety of modules and inverters to select. The number of modules in series and number of parallel strings is also specified in this step.

[1] This database has hundreds of inverters and modules and is available at the National Renewable Energy Laboratory (NREL) https://sam.nrel.gov/

4 Orientation and Output

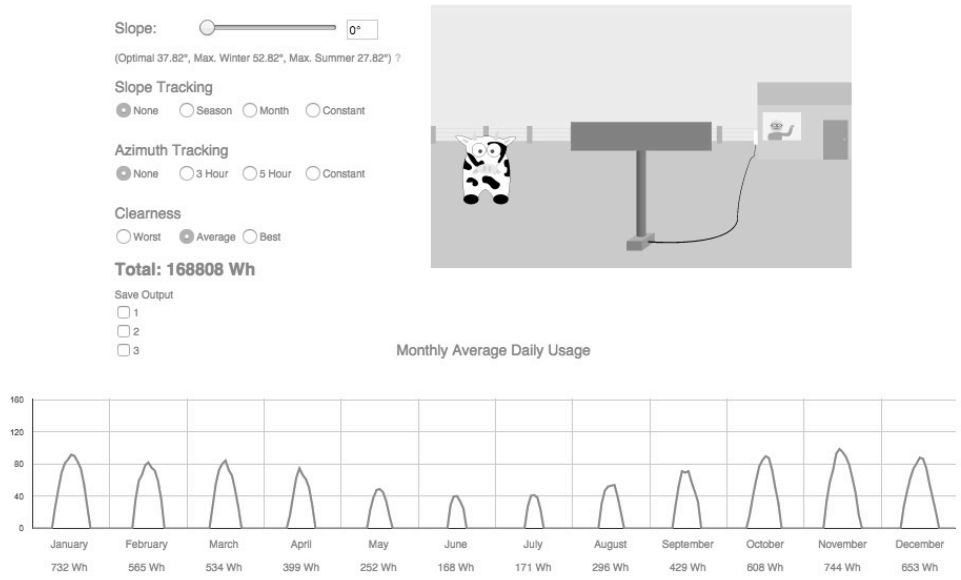


Figure 1: Last page of the tool allows the user to adjust the solar panels orientation, tracking and type of day (worst recorded day, average day and best recorded day). When adjusting these values the power output graph is updated in real time, making it visually noticeable the effects of different configurations.

The last stage of this tool is defining the module's tilt and azimuth angle. The modules tilt can be adjusted through a slider as shown in Figure 1, and the user can also choose the type of azimuth and tilt (slop) tracking. With these values known it is possible to calculate the angle of incidence between the sun's position and the module's tilt. When the angle of incidence is known it is possible to estimate the power output in four steps. The first step is to calculate the irradiance hitting the surface of the module also known as plane of array irradiance (IPOA) where POA is an acronym for plane of array. This can be calculated by adding total beam, diffuse and reflected radiation hitting the surface of the module. The second stage is using the equations proposed by Sandia Laboratories (Sandia National Laboratories and PVPMC, 2014) to estimate the cell and modules temperature, with these temperature and radiation values known we can use the Sandia PV Array Performance Model (SAPM) (King, Boyson, & Kratochvil, Photovoltaic Array Performance Model , 2004) to estimate the maximum power output of the solar panels. The final step is calculating the inverter's conversion from DC to AC power. This simulation tool uses the Sandia Performance Model for Grid-Connected Photovoltaic Inverters (King, Gonzalez, Galbraith, & Boyson, 2007). This tool also assumes that the inverter has maximum power output tracking and is able to track perfectly the maximum power output. All these calculations are done in the background, and

once completed the average daily power output graph shown in Figure 1 is updated and presented to the user. This process happens in real time allowing the user to visually see the effects of changing the module's tilt and azimuth angle.

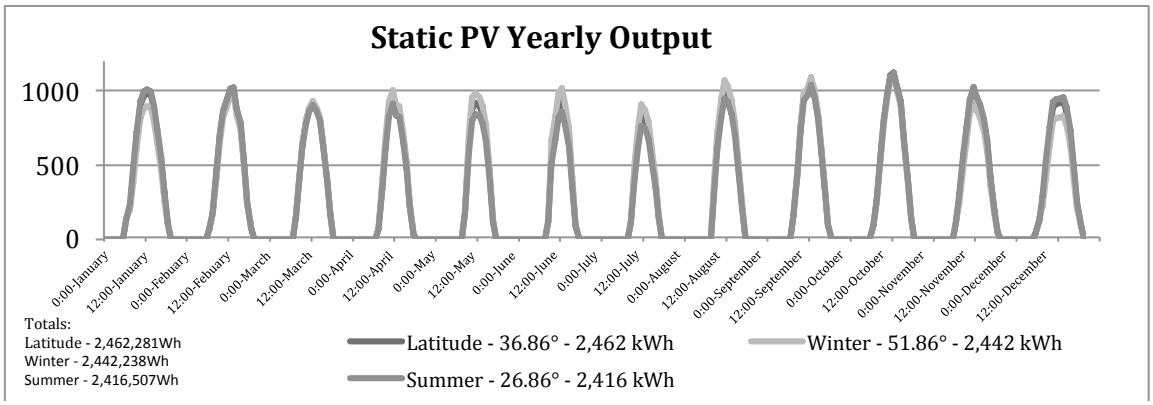


Figure 2: Graph showing different effects of solar panels' tilt in Auckland New Zealand throughout the year.

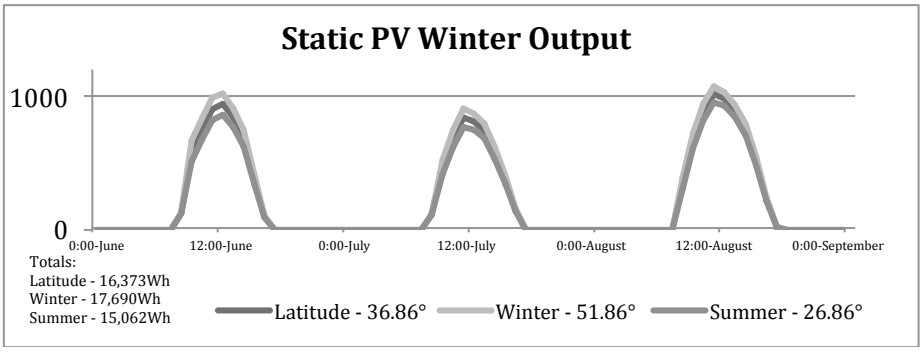


Figure 3: Graph showing different effects of solar panels' tilt throughout the winter months.

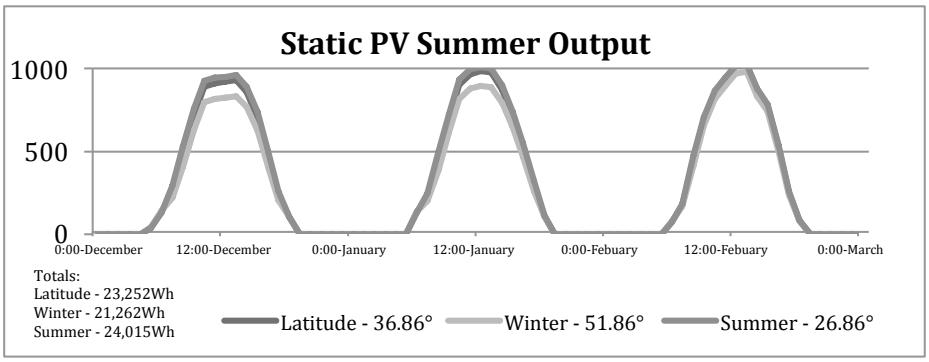


Figure 4: Graph showing different effects of solar panels' tilt throughout the summer months.

5 Results

The former sections described how this simulation tool was developed. This section is focused on looking at the results produced by this tool for different module configurations. The results are based on six 125-Watt modules in series running two in parallel making a maximum power output of 1.5 kW as shown in Table 1. The inverter is rated at a maximum of 1.5 kW, however, it is unlikely the solar panels will reach max output due to the atmospheric effects on radiation. This section describes: the effects of adjusting tilt on a static solar panel, tilt tracking, azimuth tracking and how static modules compare with single and dual axis tracking.

Module	In Series	In Parallel	Inverter	Latitude	Longitude
Mitsubishi PV-UE 125 Watt	6	2	240V HiSEL K Power 1500 Watt	-36.86	174.76

Table 1: Setup used in the tool for these results

5.1 Adjusting Tilt on a Static Solar Panel

The Authority on Sustainable Building has suggested the optimal angle for summer is -10 degrees from the angle of latitude and +15 degrees for the winter[2]. Figure 2, 3 and 4 show the tool's output with these different angles. Figure 2 shows that having a tilt equal to the latitude gives the best yearly output of 2,462 kWh, with the winter angle giving 2,442 kWh and summer with the least yearly output with only 2,416 kWh which is 46 kWh less. These values are very similar regarding yearly output, so changing the tilt between these angle values has minimal effect on yearly output with the difference between the optimal latitude angle and summer angle only being approximately 2%. However, concerning the winter months, Figure 3, the winter tilt gives the maximum output of 17,690Wh followed by latitude tilt of 16,373Wh and with the least output being the summer tilt of 15,062Wh. With these results we can see the difference between the winter tilt and latitude tilt is 8% and the difference between the winter and summer tilt being 17%. Similarly with the summer months, Figure 4, the summer tilt does produce the most output with 24,015Wh followed by the latitude tilt with 23,252Wh and the winter tilt with 21,262Wh. These results show that the summer tilt is only approximately 3% more than the latitude tilt and 13% more

[2] The Document where the Authority on Sustainable Building suggest this can be found at: www.level.org.nz

than the winter tilt. This is useful information if the user wants to get the maximum output in the winter months (e.g. they might want enough power to run the electric heater in winter) or maximum power output during the summer (e.g. having enough power for a pool pump).

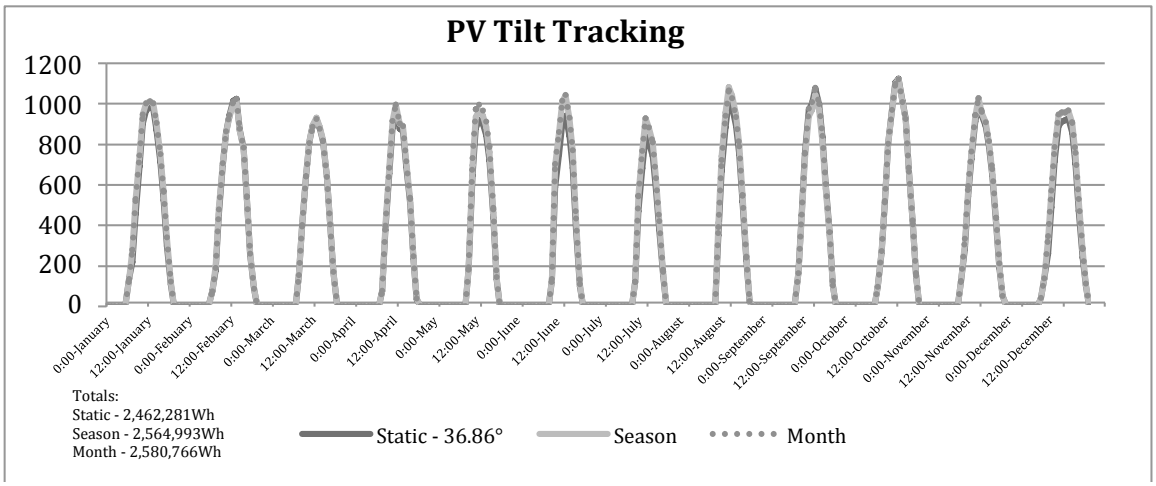


Figure 5: Graph comparing seasonal tilt tracking, monthly tilt tracking and a static PV.

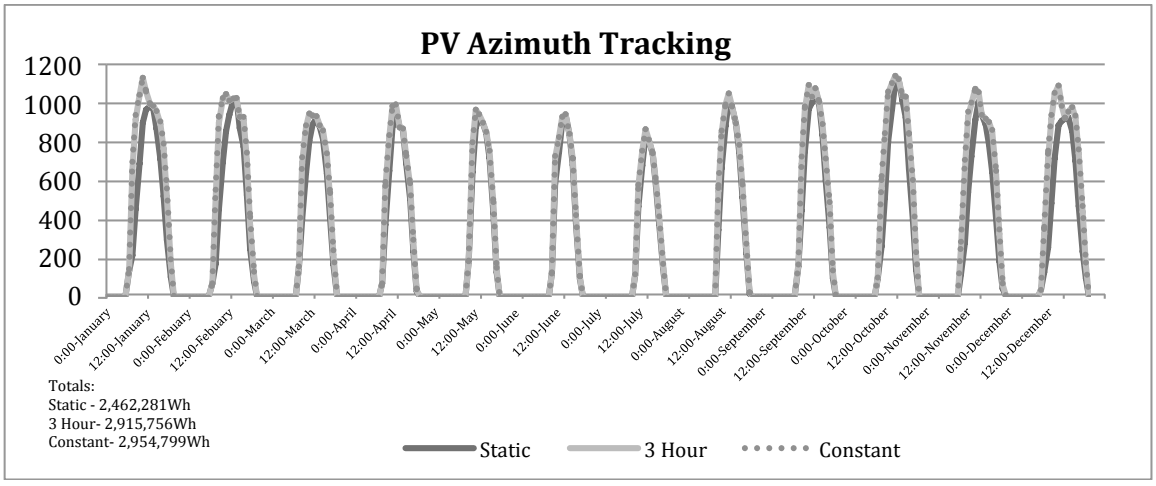


Figure 6: Graph comparing seasonal tilt tracking, monthly tilt tracking and a static PV.

5.2 Tilt Tracking

The last section shows that different solar panels at fixed tilts have a similar total yearly power output. Figure 5 shows the effects of changing the tilt every month and every season. The static fixed solar panel at a tilt of latitude gives 2,462,281Wh and changing the tilt every season would give 2,564,993Wh, which is a 4% increase. Similarly changing the tilt every month would give 2,580,766Wh, which is approximately a 5% increase. Constant tilt tracking is not added in these results because the tool simply tracks the zenith angle and does not take into consideration when the sun is behind the solar panel. This method

of constant tracking would then produce even worse results than if the solar panel was flat.

5.3 Azimuth Tracking

Azimuth tracking is more focused on following the sun as it rises from the west and sets in the east. This means it captures more sunlight hours, which is more beneficial than tracking the sun's zenith angle. In Figure 6 the static solar panel provides 2,462,281Wh and 3 hour tracking (tracking the sun's azimuth angle from 11 am to 1 pm) provides 2,915,756Wh, which is approximately an 18% increase and constant tracking provides 2,954,799Wh, which is a 20% increase. This increase is considerably larger than tilt tracking, which is why most single axis tracking, follows the sun's azimuth angle rather than its zenith angle.

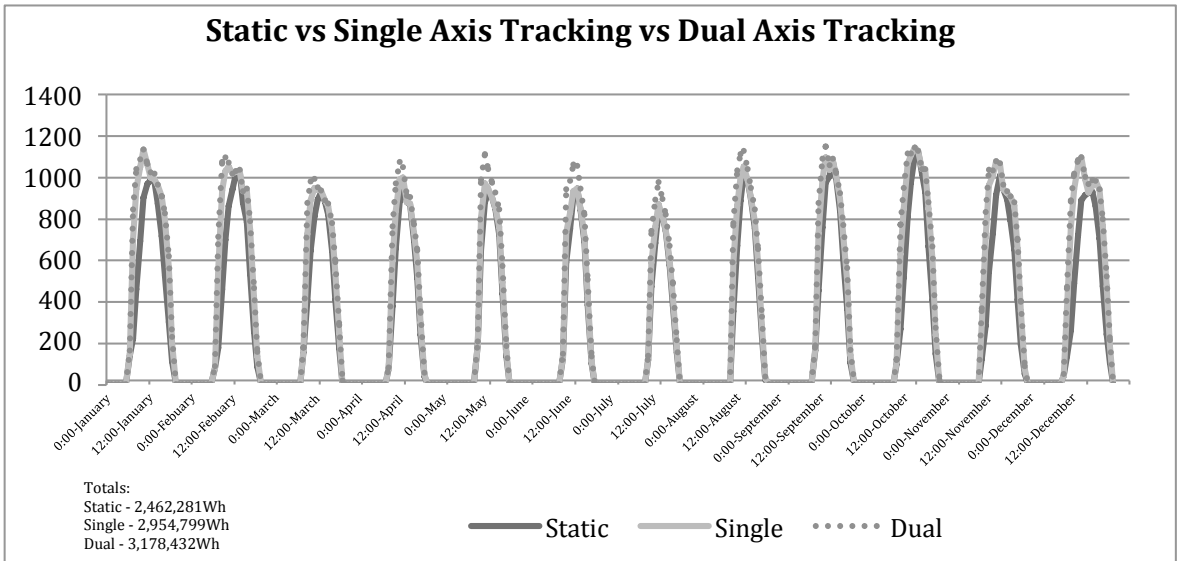


Figure 7: Graph comparing PV power output for static, single axis and dual axis tracking.

5.4 Static Tilt vs. Single Axis Tracking vs. Dual Axis Tracking

Figure 7 shows how single and dual axis tracking compare to a static angle solar panel. The single axis tracking provides an increase of 20% and dual axis provides an increase of approximately 30%. Although tracking increases the power output, it also increases the land needed and cost. The land needed for single and dual axis tracking increases to try avoid the modules casting shadows on other modules while they track the sun. The cost factor increases because of the extra parts needed initially to track the sun's position and also the ongoing cost with extra maintenance needed for these tracking parts.

5.5 Summary

It is apparent that having the solar panel at different fixed tilts can produce different power outputs at different months. If the individual wants maximum yearly power output it is suggested the tilt is as close to the latitude as possible. If maximum winter output is required, then an angle of +15 degrees on the latitude is needed. If maximum summer output is wanted, an angle of -10 degrees on the latitude is needed. Including a tracking system can increase power output, with single axis azimuth tracking giving a 20% increase and dual axis tracking with 30% increase. However, extra land and cost is needed, so further consideration by the individual is required to find out whether it is more beneficial having a tracking system or simply having more fixed solar panels. For example, dual axis tracking results in an increase of 30% but it might be more cost effective and time efficient to simply buy 30% more solar panels.

6 Conclusion

This paper describes a solar power simulation tool, which was developed to estimate a solar panel's power output at a user specified location. This tool does this in three stages. Firstly, it retrieves the location data and with this it is able to calculate the sun's position, irradiance hitting this location and estimate the wind speed and dry bulb temperature. The second stage is specifying the module and inverter configuration. The final stage is defining the orientation of the solar panel, which will simulate the power output of the solar panel(s) throughout the year. The advantage of this tool is that the user can change the orientation in real time by adjusting a slider, and immediately see the effects on power output. The user can then use these results to compare different

orientations and tracking systems to find the best solution for their power requirements.

References

[1] King, D. L., Boyson, W. E., & Kratochvil, J. A. (2004, August). *Photovoltaic Array Performance Model* . Retrieved March 2015 from Sandia National Laboratories : http://prod.sandia.gov/techlib/access-control.cgi/2004/043535.pdf

[2] King, D. L., Gonzalez, S., Galbraith, G. M., & Boyson, W. E. (2007, September). *Performance Model for Grid-Connected Photovoltaic Inverters* . Retrieved March 2015 from Sandia National Laboratories: http://prod.sandia.gov/techlib/access-control.cgi/2007/075036.pdf

[3] Laue, E. (1970). The measurement of solar spectral irradiance at different terrestrial elevations. *Solar Energy , 13* (1), 43–50.

[4] Maxwell, E. L. (1987). *A Quasi-Physical Model for Converting Hourly Global Horizontal to Direct Normal Insolation*. Golden, CO, Solar Energy Research Institute.

[5] Sandia National Laboratories and PVPMC. (2014). *Modeling Steps*. Retrieved February 2015 from PV Performance Modeling Collaborative: https://pvpmc.sandia.gov

Author Index

Alhonsuo, Mira, 381
Al-Jaafreh, Abdel-karim, 323
AlTarawneh, Ragaad, 293
AlTarawneh, Ragaad, 323
André, Elisabeth, 565, 585
Apperley, Mark, 623, 633
Archdeacon, John, 81
Ardito, Carmelo, 239, 497
Avouris, Nikolaos, 465

Barbosa, Simone D.J., 211
Baroffio, Giorgia, 555
Barricelli, Barbara Rita, 129
Behrendt, Marcus, 119
Bellino, Alessio, 3
Bernhaupt, Regina, 203, 259
Billestrup, Jane, 407
Billman, Dorrit, 81
Bobeth, Jan, 187
Boll, Susanne, 301
Boscarol, Maurizio, 497
Bowen, Judy, 137
Bruun, Anders, 407
Buchanan, George, 95
Buchner, Roland, 145
Burmester, Michael, 153
Burrows, Rachel, 547

Cajander, Åsa, 161, 249, 357
Campos, Pedro, 179
Castelletti, Andrea, 555
Chung, Cindy K., 293
Clemmensen, Torkil, 171
Colley, Ashley, 381
Coninx, Karin, 345
Convertino, Gregorio, 449
Cordeiro de Paula, Felipe, 211
Costabile, Maria Francesca, 239
Costanza, Enrico, 539
Cunningham, Sally Jo, 137

Darin, Ticianne, 61
Day, Jonathan, 95
Decancq, Jasmien, 345
Dermody, Fiona, 49
Deshmukh, Rohit, 81
Dobbels, Fabienne, 399

Elprama, Shirley, 345
Evans, W. Douglas, 369

Fahssi, Racim, 101
Farren, Margaret, 49
Fayollas, Camille, 101
Feary, Michael, 81
Feuerstack, Sebastian, 119
Fidas, Christos, 465
Fischer, Holger, 215
Forbrig, Peter, 223
Fraternali, Piero, 555
Fuchsberger, Verena, 145
Furci, Ferdinando, 507
Fussel, Susan R., 27

Garg, Anant, 179
Gattol, Valentin, 187
Geerts, David, 195, 399
Ghosh, Sanjay, 43
Gonçalves, Frederica, 179
Goyal, Nitesh, 27
Greis, Miriam, 15
Grünloh, Christiane, 357

Hak, Jean-Luc, 267
Häkkilä, Jonna, 381
Hammer, Stephan, 565
Harkke, Ville, 513
Heintz, Matthias, 485
Herczeg, Michael, 223
Hermosa Perrino, Cristina, 153
Heuten, Wilko, 301
Hinze, Annika, 137
Holbrook, Jon, 81
Holl, Konstantin, 313

Jacobs, An, 345
Johansson, Stefan, 39
Johnson, Hilary, 547
Johnson, Peter, 547
Jorge, Clinton, 605

Killeen, Chris, 595
Kuusinen, Kati, 233
Kyfonidis, Charalampos, 387

Laib, Magdalena, 153
Lanzilotti, Rosa, 239, 497
Larusdottir, Marta, 161, 249
Law, Effie, 249, 485
Lehner, Ulrich, 187
Lind, Thomas, 161
Lugrin, Birgit, 565
Luxton-Reilly, Andrew, 9

Makri, Stephann, 95
Manciet, François, 203, 259
Marshall, Carl S., 293
Martinie, Célia, 101
Masoodian, Masood, 585
Mastrangelo, Simon, 497
Mbaya, Ibrahim R., 73
McGee-Lennon, Marilyn, 387
Meneweger, Thomas, 335
Micheel, Isabel, 555
Mnkandla, Ernest, 67
Mori, Giulio, 507
Müller, Christian, 313
Müller, Jens, 615

Nass, Claudia, 313
Nauwerck, Gerolf, 161
Nisi, Valentina, 575, 605
Novak, Jasminko, 555
Nunes, Nuno J., 575, 605

Obrist, Marianna, 335
Olsson, Thomas, 521
Orvieto, Ivan, 465
Ozoh, Patrick, 623

Palanque, Philippe, 101
Panneer, Selvakumar, 293
Paternò, Fabio, 507
Pereira, Lucas, 575
Pirker, Michael, 203
Pirker, Michael M., 259
Plimmer, Beryl, 9

Queßeleit, Marie-Luise, 153
Quintal, Filipe, 605

Rantakari, Juho, 381
Regal, Georg, 187
Rexhepi4, Hanife, 357
Rist, Thomas, 615
Rizzoli, Andrea-Emilio, 555
Rogers, Alex, 539
Roto, Virpi, 521
Rukh Humayoun, Shah, 323

Sauer, Stefan, 215
Savioja, Paula, 521
Schiotka, Alexander, 301
Segmüller, Fabian, 565
Senekal, Marjanne, 369
Silva, Thiago R., 267
Stage, Jan, 407
Stewart, Michael, 81
Steyn, Nelia P., 369
Suppers, Joris, 633
Sutherland, Alistair, 49
Sutherland, Craig, 9
Swallow, David, 55

Tai Dang, Chi, 585
Tarkkanen, Kimmo, 513
Timmermann, Janko, 301
Toko, Guy, 67
Tolias, Evangelos, 539
Tscheligi, Manfred, 145, 187, 335

Väätäjä, Heli, 521
Valtolina, Stefano, 129
Van den Bergh, Jan, 345

van der Veer, Gerrit C., 283
Vandenberghe, Bert, 195, 399
Vanhoof, Jasper, 399
Varsaluoma, Jari, 521
Vasiliou, Christina, 33
Virtanen, Lasse, 381

Weiss, Astrid, 145
Winckler, Marco, 267
Wortelen, Bertram, 119
Wurhofer, Daniela, 21

Yigitbas, Enes, 215
Yuan, Xiaojun, 529

Zeiner, Katharina Maria, 153

University of Bamberg Press

INTERACT is among the world's top conferences in Human-Computer Interaction. Starting with the first INTERACT conference in 1990, this conference series has been organised under the aegis of the Technical Committee 13 on Human-Computer Interaction of the UNESCO International Federation for Information Processing (IFIP). This committee aims at developing the science and technology of the interaction between humans and computing devices.

The 15th IFIP TC.13 International Conference on Human-Computer Interaction - INTERACT 2015 took place from 14 to 18 September 2015 in Bamberg, Germany. The theme of INTERACT 2015 was „Connection.Tradition.Innovation". This volume presents the Adjunct Proceedings - it contains the position papers for the students of the Doctoral Consortium as well as the position papers of the participants of the various workshops.

eISBN: 978-3-86309-353-2

www.uni-bamberg.de/ubp