Masaaki Kurosu (Ed.)

Human-Computer Interaction

Users and Contexts of Use

15th International Conference, HCI International 2013
Las Vegas, NV, USA, July 2013
Proceedings, Part III

LNCS 8006

3 Part III

Lecture Notes in Computer Science 8006

Commenced Publication in 1973
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Masaaki Kurosu (Ed.)

Human-Computer Interaction

Users and Contexts of Use

15th International Conference, HCI International 2013
Las Vegas, NV, USA, July 21-26, 2013
Proceedings, Part III

Springer

Volume Editor

Masaaki Kurosu
The Open University of Japan
2-11 Wakaba, Mihama-ku, Chiba-shi 261-8586, Japan
E-mail: masaakikurosu@spa.nifty.com

ISSN 0302-9743 e-ISSN 1611-3349
ISBN 978-3-642-39264-1 e-ISBN 978-3-642-39265-8
DOI 10.1007/978-3-642-39265-8
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2013941394

CR Subject Classification (1998): H.5, H.3, H.4, K.3, K.6

LNCS Sublibrary: SL 3 – Information Systems and Application, incl. Internet/Web and HCI

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Foreword

The 15th International Conference on Human–Computer Interaction, HCI International 2013, was held in Las Vegas, Nevada, USA, 21–26 July 2013, incorporating 12 conferences / thematic areas:

Thematic areas:

- Human–Computer Interaction
- Human Interface and the Management of Information

Affiliated conferences:

- 10th International Conference on Engineering Psychology and Cognitive Ergonomics
- 7th International Conference on Universal Access in Human–Computer Interaction
- 5th International Conference on Virtual, Augmented and Mixed Reality
- 5th International Conference on Cross-Cultural Design
- 5th International Conference on Online Communities and Social Computing
- 7th International Conference on Augmented Cognition
- 4th International Conference on Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management
- 2nd International Conference on Design, User Experience and Usability
- 1st International Conference on Distributed, Ambient and Pervasive Interactions
- 1st International Conference on Human Aspects of Information Security, Privacy and Trust

A total of 5210 individuals from academia, research institutes, industry and governmental agencies from 70 countries submitted contributions, and 1666 papers and 303 posters were included in the program. These papers address the latest research and development efforts and highlight the human aspects of design and use of computing systems. The papers accepted for presentation thoroughly cover the entire field of Human–Computer Interaction, addressing major advances in knowledge and effective use of computers in a variety of application areas.

This volume, edited by Masaaki Kurosu, contains papers focusing on the thematic area of Human–Computer Interaction, and addressing the following major topics:

- Identity, Privacy and Trust
- User Studies
- Interaction for Society and Community
- HCI for Business and Innovation

The remaining volumes of the HCI International 2013 proceedings are:

- Volume 1, LNCS 8004, Human–Computer Interaction: Human-Centred Design Approaches, Methods, Tools and Environments (Part I), edited by Masaaki Kurosu
- Volume 2, LNCS 8005, Human–Computer Interaction: Applications and Services (Part II), edited by Masaaki Kurosu
- Volume 4, LNCS 8007, Human–Computer Interaction: Interaction Modalities and Techniques (Part IV), edited by Masaaki Kurosu
- Volume 5, LNCS 8008, Human–Computer Interaction: Towards Intelligent and Implicit Interaction (Part V), edited by Masaaki Kurosu
- Volume 6, LNCS 8009, Universal Access in Human–Computer Interaction: Design Methods, Tools and Interaction Techniques for eInclusion (Part I), edited by Constantine Stephanidis and Margherita Antona
- Volume 7, LNCS 8010, Universal Access in Human–Computer Interaction: User and Context Diversity (Part II), edited by Constantine Stephanidis and Margherita Antona
- Volume 8, LNCS 8011, Universal Access in Human–Computer Interaction: Applications and Services for Quality of Life (Part III), edited by Constantine Stephanidis and Margherita Antona
- Volume 9, LNCS 8012, Design, User Experience, and Usability: Design Philosophy, Methods and Tools (Part I), edited by Aaron Marcus
- Volume 10, LNCS 8013, Design, User Experience, and Usability: Health, Learning, Playing, Cultural, and Cross-Cultural User Experience (Part II), edited by Aaron Marcus
- Volume 11, LNCS 8014, Design, User Experience, and Usability: User Experience in Novel Technological Environments (Part III), edited by Aaron Marcus
- Volume 12, LNCS 8015, Design, User Experience, and Usability: Web, Mobile and Product Design (Part IV), edited by Aaron Marcus
- Volume 13, LNCS 8016, Human Interface and the Management of Information: Information and Interaction Design (Part I), edited by Sakae Yamamoto
- Volume 14, LNCS 8017, Human Interface and the Management of Information: Information and Interaction for Health, Safety, Mobility and Complex Environments (Part II), edited by Sakae Yamamoto
- Volume 15, LNCS 8018, Human Interface and the Management of Information: Information and Interaction for Learning, Culture, Collaboration and Business (Part III), edited by Sakae Yamamoto
- Volume 16, LNAI 8019, Engineering Psychology and Cognitive Ergonomics: Understanding Human Cognition (Part I), edited by Don Harris
- Volume 17, LNAI 8020, Engineering Psychology and Cognitive Ergonomics: Applications and Services (Part II), edited by Don Harris
- Volume 18, LNCS 8021, Virtual, Augmented and Mixed Reality: Designing and Developing Augmented and Virtual Environments (Part I), edited by Randall Shumaker
- Volume 19, LNCS 8022, Virtual, Augmented and Mixed Reality: Systems and Applications (Part II), edited by Randall Shumaker

- Volume 20, LNCS 8023, Cross-Cultural Design: Methods, Practice and Case Studies (Part I), edited by P.L. Patrick Rau
- Volume 21, LNCS 8024, Cross-Cultural Design: Cultural Differences in Everyday Life (Part II), edited by P.L. Patrick Rau
- Volume 22, LNCS 8025, Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management: Healthcare and Safety of the Environment and Transport (Part I), edited by Vincent G. Duffy
- Volume 23, LNCS 8026, Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management: Human Body Modeling and Ergonomics (Part II), edited by Vincent G. Duffy
- Volume 24, LNAI 8027, Foundations of Augmented Cognition, edited by Dylan D. Schmorrow and Cali M. Fidopiastis
- Volume 25, LNCS 8028, Distributed, Ambient and Pervasive Interactions, edited by Norbert Streitz and Constantine Stephanidis
- Volume 26, LNCS 8029, Online Communities and Social Computing, edited by A. Ant Ozok and Panayiotis Zaphiris
- Volume 27, LNCS 8030, Human Aspects of Information Security, Privacy and Trust, edited by Louis Marinos and Ioannis Askoxylakis
- Volume 28, CCIS 373, HCI International 2013 Posters Proceedings (Part I), edited by Constantine Stephanidis
- Volume 29, CCIS 374, HCI International 2013 Posters Proceedings (Part II), edited by Constantine Stephanidis

I would like to thank the Program Chairs and the members of the Program Boards of all affiliated conferences and thematic areas, listed below, for their contribution to the highest scientific quality and the overall success of the HCI International 2013 conference.

This conference could not have been possible without the continuous support and advice of the Founding Chair and Conference Scientific Advisor, Prof. Gavriel Salvendy, as well as the dedicated work and outstanding efforts of the Communications Chair and Editor of HCI International News, Abbas Moallem.

I would also like to thank for their contribution towards the smooth organization of the HCI International 2013 Conference the members of the Human–Computer Interaction Laboratory of ICS-FORTH, and in particular George Paparoulis, Maria Pitsoulaki, Stavroula Ntoa, Maria Bouhli and George Kapnas.

May 2013

Constantine Stephanidis
General Chair, HCI International 2013

Organization

Human–Computer Interaction

Program Chair: Masaaki Kurosu, Japan

Jose Abdelnour-Nocera, UK
Sebastiano Bagnara, Italy
Simone Barbosa, Brazil
Tomas Berns, Sweden
Nigel Bevan, UK
Simone Borsci, UK
Apala Lahiri Chavan, India
Sherry Chen, Taiwan
Kevin Clark, USA
Torkil Clemmensen, Denmark
Xiaowen Fang, USA
Shin'ichi Fukuzumi, Japan
Vicki Hanson, UK
Ayako Hashizume, Japan
Anzai Hiroyuki, Italy
Sheue-Ling Hwang, Taiwan
Wonil Hwang, South Korea
Minna Isomursu, Finland
Yong Gu Ji, South Korea
Esther Jun, USA
Mitsuhiko Karashima, Japan
Kyungdoh Kim, South Korea
Heidi Krömker, Germany
Chen Ling, USA
Yan Liu, USA
Zhengjie Liu, P.R. China
Loïc Martínez Normand, Spain
Chang S. Nam, USA
Naoko Okuizumi, Japan
Noriko Osaka, Japan
Philippe Palanque, France
Hans Persson, Sweden
Ling Rothrock, USA
Naoki Sakakibara, Japan
Dominique Scapin, France
Guangfeng Song, USA
Sanjay Tripathi, India
Chui Yin Wong, Malaysia
Toshiki Yamaoka, Japan
Kazuhiko Yamazaki, Japan
Ryoji Yoshitake, Japan
Silvia Zimmermann, Switzerland

Human Interface and the Management of Information

Program Chair: Sakae Yamamoto, Japan

Hans-Jorg Bullinger, Germany
Alan Chan, Hong Kong
Gilsoo Cho, South Korea
Jon R. Gunderson, USA
Shin'ichi Fukuzumi, Japan
Michitaka Hirose, Japan
Jhilmil Jain, USA
Yasufumi Kume, Japan
Mark Lehto, USA
Hiroyuki Miki, Japan
Hirohiko Mori, Japan
Fiona Fui-Hoon Nah, USA
Shogo Nishida, Japan
Robert Proctor, USA
Youngho Rhee, South Korea
Katsunori Shimohara, Japan

Michale Smith, USA
Tsutomu Tabe, Japan
Hiroshi Tsuji, Japan
Kim-Phuong Vu, USA
Tomio Watanabe, Japan
Hidekazu Yoshikawa, Japan

Engineering Psychology and Cognitive Ergonomics

Program Chair: Don Harris, UK

Guy Andre Boy, USA
Joakim Dahlman, Sweden
Trevor Dobbins, UK
Mike Feary, USA
Shan Fu, P.R. China
Michaela Heese, Austria
Hung-Sying Jing, Taiwan
Wen-Chin Li, Taiwan
Mark A. Neerincx, The Netherlands
Jan M. Noyes, UK
Taezoon Park, Singapore
Paul Salmon, Australia
Axel Schulte, Germany
Siraj Shaikh, UK
Sarah C. Sharples, UK
Anthony Smoker, UK
Neville A. Stanton, UK
Alex Stedmon, UK
Xianghong Sun, P.R. China
Andrew Thatcher, South Africa
Matthew J.W. Thomas, Australia
Rolf Zon, The Netherlands

Universal Access in Human–Computer Interaction

Program Chairs: Constantine Stephanidis, Greece, and Margherita Antona, Greece

Julio Abascal, Spain
Ray Adams, UK
Gisela Susanne Bahr, USA
Margit Betke, USA
Christian Bühler, Germany
Stefan Carmien, Spain
Jerzy Charytonowicz, Poland
Carlos Duarte, Portugal
Pier Luigi Emiliani, Italy
Qin Gao, P.R. China
Andrina Granić, Croatia
Andreas Holzinger, Austria
Josette Jones, USA
Simeon Keates, UK
Georgios Kouroupetroglou, Greece
Patrick Langdon, UK
Seongil Lee, Korea
Ana Isabel B.B. Paraguay, Brazil
Helen Petrie, UK
Michael Pieper, Germany
Enrico Pontelli, USA
Jaime Sanchez, Chile
Anthony Savidis, Greece
Christian Stary, Austria
Hirotada Ueda, Japan
Gerhard Weber, Germany
Harald Weber, Germany

Virtual, Augmented and Mixed Reality

Program Chair: Randall Shumaker, USA

Waymon Armstrong, USA
Juan Cendan, USA
Rudy Darken, USA
Cali M. Fidopiastis, USA
Charles Hughes, USA
David Kaber, USA
Hirokazu Kato, Japan
Denis Laurendeau, Canada
Fotis Liarokapis, UK
Mark Livingston, USA
Michael Macedonia, USA
Gordon Mair, UK
Jose San Martin, Spain
Jacquelyn Morie, USA
Albert "Skip" Rizzo, USA
Kay Stanney, USA
Christopher Stapleton, USA
Gregory Welch, USA

Cross-Cultural Design

Program Chair: P.L. Patrick Rau, P.R. China

Pilsung Choe, P.R. China
Henry Been-Lirn Duh, Singapore
Vanessa Evers, The Netherlands
Paul Fu, USA
Zhiyong Fu, P.R. China
Fu Guo, P.R. China
Sung H. Han, Korea
Toshikazu Kato, Japan
Dyi-Yih Michael Lin, Taiwan
Rungtai Lin, Taiwan
Sheau-Farn Max Liang, Taiwan
Liang Ma, P.R. China
Alexander Mädche, Germany
Katsuhiko Ogawa, Japan
Tom Plocher, USA
Kerstin Röse, Germany
Supriya Singh, Australia
Hsiu-Ping Yueh, Taiwan
Liang (Leon) Zeng, USA
Chen Zhao, USA

Online Communities and Social Computing

Program Chairs: A. Ant Ozok, USA, and Panayiotis Zaphiris, Cyprus

Areej Al-Wabil, Saudi Arabia
Leonelo Almeida, Brazil
Bjørn Andersen, Norway
Chee Siang Ang, UK
Aneesha Bakharia, Australia
Ania Bobrowicz, UK
Paul Cairns, UK
Farzin Deravi, UK
Andri Ioannou, Cyprus
Slava Kisilevich, Germany
Niki Lambropoulos, Greece
Effie Law, Switzerland
Soo Ling Lim, UK
Fernando Loizides, Cyprus
Gabriele Meiselwitz, USA
Anthony Norcio, USA
Elaine Raybourn, USA
Panote Siriaraya, UK
David Stuart, UK
June Wei, USA

Augmented Cognition

Program Chairs: Dylan D. Schmorrow, USA, and Cali M. Fidopiastis, USA

Robert Arrabito, Canada
Richard Backs, USA
Chris Berka, USA
Joseph Cohn, USA
Martha E. Crosby, USA
Julie Drexler, USA
Ivy Estabrooke, USA
Chris Forsythe, USA
Wai Tat Fu, USA
Rodolphe Gentili, USA
Marc Grootjen, The Netherlands
Jefferson Grubb, USA
Ming Hou, Canada
Santosh Mathan, USA
Rob Matthews, Australia
Dennis McBride, USA
Jeff Morrison, USA
Mark A. Neerincx, The Netherlands
Denise Nicholson, USA
Banu Onaral, USA
Lee Sciarini, USA
Kay Stanney, USA
Roy Stripling, USA
Rob Taylor, UK
Karl van Orden, USA

Digital Human Modeling and Applications in Health, Safety, Ergonomics and Risk Management

Program Chair: Vincent G. Duffy, USA and Russia

Karim Abdel-Malek, USA
Giuseppe Andreoni, Italy
Daniel Carruth, USA
Eliza Yingzi Du, USA
Enda Fallon, Ireland
Afzal Godil, USA
Ravindra Goonetilleke, Hong Kong
Bo Hoege, Germany
Waldemar Karwowski, USA
Zhizhong Li, P.R. China
Kang Li, USA
Tim Marler, USA
Michelle Robertson, USA
Matthias Rötting, Germany
Peter Vink, The Netherlands
Mao-Jiun Wang, Taiwan
Xuguang Wang, France
Jingzhou (James) Yang, USA
Xiugan Yuan, P.R. China
Gülcin Yücel Hoge, Germany

Design, User Experience, and Usability

Program Chair: Aaron Marcus, USA

Sisira Adikari, Australia
Ronald Baecker, Canada
Arne Berger, Germany
Jamie Blustein, Canada
Ana Boa-Ventura, USA
Jan Brejcha, Czech Republic
Lorenzo Cantoni, Switzerland
Maximilian Eibl, Germany

Anthony Faiola, USA
Emilie Gould, USA
Zelda Harrison, USA
Rüdiger Heimgärtner, Germany
Brigitte Herrmann, Germany
Steffen Hess, Germany
Kaleem Khan, Canada
Jennifer McGinn, USA
Francisco Rebelo, Portugal
Michael Renner, Switzerland
Kerem Rızvanoğlu, Turkey
Marcelo Soares, Brazil
Christian Sturm, Germany
Michele Visciola, Italy

Distributed, Ambient and Pervasive Interactions

Program Chairs: Norbert Streitz, Germany, and Constantine Stephanidis, Greece

Emile Aarts, The Netherlands
Adnan Abu-Dayya, Qatar
Juan Carlos Augusto, UK
Boris de Ruyter, The Netherlands
Anind Dey, USA
Dimitris Grammenos, Greece
Nuno M. Guimaraes, Portugal
Shin'ichi Konomi, Japan
Carsten Magerkurth, Switzerland
Christian Müller-Tomfelde, Australia
Fabio Paternó, Italy
Gilles Privat, France
Harald Reiterer, Germany
Carsten Röcker, Germany
Reiner Wichert, Germany
Woontack Woo, South Korea
Xenophon Zabulis, Greece

Human Aspects of Information Security, Privacy and Trust

Program Chairs: Louis Marinos, ENISA EU, and Ioannis Askoxylakis, Greece

Claudio Agostino Ardagna, Italy
Zinaida Benenson, Germany
Daniele Catteddu, Italy
Raoul Chiesa, Italy
Bryan Cline, USA
Sadie Creese, UK
Jorge Cuellar, Germany
Marc Dacier, USA
Dieter Gollmann, Germany
Kirstie Hawkey, Canada
Jaap-Henk Hoepman, The Netherlands
Cagatay Karabat, Turkey
Angelos Keromytis, USA
Ayako Komatsu, Japan
Ronald Leenes, The Netherlands
Javier Lopez, Spain
Steve Marsh, Canada
Gregorio Martinez, Spain
Emilio Mordini, Italy
Yuko Murayama, Japan
Masakatsu Nishigaki, Japan
Aljosa Pasic, Spain
Milan Petković, The Netherlands
Joachim Posegga, Germany
Jean-Jacques Quisquater, Belgium
Damien Sauveron, France
George Spanoudakis, UK
Kerry-Lynn Thomson, South Africa

Julien Touzeau, France
Theo Tryfonas, UK
João Vilela, Portugal
Claire Vishik, UK
Melanie Volkamer, Germany

External Reviewers

Maysoon Abulkhair, Saudi Arabia
Ilia Adami, Greece
Vishal Barot, UK
Stephan Böhm, Germany
Vassilis Charissis, UK
Francisco Cipolla-Ficarra, Spain
Maria De Marsico, Italy
Marc Fabri, UK
David Fonseca, Spain
Linda Harley, USA
Yasushi Ikei, Japan
Wei Ji, USA
Nouf Khashman, Canada
John Killilea, USA
Iosif Klironomos, Greece
Ute Klotz, Switzerland
Maria Korozi, Greece
Kentaro Kotani, Japan
Vassilis Kouroumalis, Greece
Stephanie Lackey, USA
Janelle LaMarche, USA
Asterios Leonidis, Greece
Nickolas Macchiarella, USA
George Margetis, Greece
Matthew Marraffino, USA
Joseph Mercado, USA
Claudia Mont'Alvão, Brazil
Yoichi Motomura, Japan
Karsten Nebe, Germany
Stavroula Ntoa, Greece
Martin Osen, Austria
Stephen Prior, UK
Farid Shirazi, Canada
Jan Stelovsky, USA
Sarah Swierenga, USA

HCI International 2014

The 16th International Conference on Human–Computer Interaction, HCI International 2014, will be held jointly with the affiliated conferences in the summer of 2014. It will cover a broad spectrum of themes related to Human–Computer Interaction, including theoretical issues, methods, tools, processes and case studies in HCI design, as well as novel interaction techniques, interfaces and applications. The proceedings will be published by Springer. More information about the topics, as well as the venue and dates of the conference, will be announced through the HCI International Conference series website: `http://www.hci-international.org/`

General Chair
Professor Constantine Stephanidis
University of Crete and ICS-FORTH
Heraklion, Crete, Greece
Email: `cs@ics.forth.gr`

Table of Contents – Part III

Identity, Privacy and Trust

SP-CIP: A Secure and Privacy Aware Patient Centric Information Portal 3
Subrata Acharya

Identity Management through "Profiles": Prototyping an Online Information Segregation Service 10
Julio Angulo and Erik Wästlund

An Empirical Investigation of Similarity-Driven Trust Dynamics in a Social Network 20
Yugo Hayashi, Victor Kryssanov, and Hitoshi Ogawa

The di.me User Interface: Concepts for Sharing Personal Information via Multiple Identities in a Decentralized Social Network 29
Fabian Hermann, Andreas Schuller, Simon Thiel, Christian Knecht, and Simon Scerri

Trust and Privacy in the di.me Userware 39
Marcel Heupel, Mohamed Bourimi, and Doğan Kesdoğan

From Living Space to Urban Quarter: Acceptance of ICT Monitoring Solutions in an Ageing Society 49
Simon Himmel, Martina Ziefle, and Katrin Arning

The Impact of Explanation Dialogues on Human-Computer Trust 59
Florian Nothdurft, Tobias Heinroth, and Wolfgang Minker

Trustworthy and Inclusive Identity Management for Applications in Social Media 68
Till Halbach Røssvoll and Lothar Fritsch

Towards a Design Guideline of Visual Cryptography on Stereoscopic Displays 78
Shih-Lung Tsai and Chao-Hua Wen

User Studies

Interface Design for Minimizing Loss of Context in In-situ Remote Robot Control 87
Jong-gil Ahn and Gerard Jounghyun Kim

Assessing the Impact of Latency and Jitter on the Perceived Quality of Call of Duty Modern Warfare 2 97
Rahul Amin, France Jackson, Juan Gilbert, Jim Martin, and Terry Shaw

The Effects of Negative Interaction Feedback in a Web Navigation Assistant 107
Marcelo G. Armentano and Analía A. Amandi

Automatic Layout Generation for Digital Photo Albums: A User Study 117
Francine B. Bergmann, Isabel H. Manssour, Milene S. Silveira, and João Batista S. de Oliveira

Icon Design for Older Users of Project Management Software 127
Christina Bröhl, Jennifer Bützler, Nicole Jochems, and Christopher M. Schlick

Search Strategies in Hypermedia Navigation and Spatial Abilities: A Comparison with Physical Navigation 136
Aurélie Brouwers

Data Visualisation, User Experience and Context: A Case Study from Fantasy Sport 146
Rob Euman and José Abdelnour-Nocera

A Pedestrian Navigation Method for User's Safe and Easy Wayfinding 156
Hiroshi Furukawa and Yutaka Nakamura

The Effect of Information Quantity on Cbp Interface in the Advanced Nuclear Power Plant 166
Min-Chih Hsieh and Sheue-Ling Hwang

Development of Communication Support System at Mealtimes Using Tabletop Interface 174
Junko Itou, Shizuka Takao, and Jun Munemori

Improving User Performance in Conditional Probability Problems with Computer-Generated Diagrams 183
Vince Kellen, Susy Chan, and Xiaowen Fang

Establishing a Cognitive Map of Public Place for Blind and Visual Impaired by Using IVEO Hands-On Learning System 193
Qing-Wen Lin, Sheue-Ling Hwang, and Jan-Li Wang

An Application of the Ballistic Movement Method for Evaluating Computer Mice 199
Ray F. Lin, Ching-Wen Chung, Yi-Chien Tsai, and Chi-Yu Huang

AugmentedBacklight: Expansion of LCD Backlights Using Lighting Methods in the Real World 209
Maho Oki, Koji Tsukada, and Itiro Siio

An Exploratory Study to Understand Knowledge-Sharing in Data-Intensive Science 217
Jongsoon Park and Joseph L. Gabbard

Interaction of the Elderly Viewer with Additional Multimedia Content to Support the Appreciation of Television Programs 227
Kamila Rodrigues, Vânia Paula de Almeida Neris, and Cesar Augusto Camillo Teixeira

Exploring Children's Attitudes towards Static and Moving Humanoid Robots 237
Fang-Wu Tung and Tsen-Yao Chang

Interaction for Society and Community

Mobile Money Services in Uganda: Design Gaps and Recommendations 249
Rehema Baguma

A Static and Dynamic Recommendations System for Best Practice Networks 259
Pierfrancesco Bellini, Ivan Bruno, Paolo Nesi, and Michela Paolucci

Connecting Electric Vehicles and Green Energy 269
Peter Dempster

SOCIETY: A Social Reading Application to Join Education and Social Network Experience 277
Elena Guercio, Fabio Luciano Mondin, Maurizio Belluati, and Lucia Longo

The Evaluation of a Voting Web Based Application 285
Linda Harley, Keith Kline, Jerry Ray, Carrie Bell, Andrew Baranak, Chandler Price, Matthew Hung, and Brad Fain

A Refuge Location Prediction System for When a Tsunami Has Occurred 295
Akhiro Kawabe, Tomoko Izumi, and Yoshio Nakatani

Further Benefit of a Kind of Inconvenience for Social Information Systems 301
Hiroshi Kawakami

The Present Condition and Problems for Elderly People Participating in Communities 307
Noriyuki Kobayashi, Shohei Yamane, Kotaro Ohori, and Akihiko Obata

Applying to Twitter Networks of a Community Extraction Method Using Intersection Graph and Semantic Analysis 314
Toshiya Kuramochi, Naoki Okada, Kyohei Tanikawa, Yoshinori Hijikata, and Shogo Nishida

Search Engine Accessibility for Low-Literate Users 324
Débora Maurmo Modesto, Simone Bacellar Leal Ferreira, and Aline Silva Alves

Identifying and Representing Elements of Local Contexts in Namibia . . . 332
Kasper Rodil, Kasper Løvborg Jensen, Matthias Rehm, and Heike Winschiers-Theophilus

A Framework for Community-Oriented Mobile Interaction Design in Emerging Regions 342
Monica Sebillo, Genoveffa Tortora, Giuliana Vitiello, Pasquale Di Giovanni, and Marco Romano

The Effects of (Social) Media on Revolutions – Perspectives from Egypt and the Arab Spring 352
Christian Sturm and Hossam Amer

Memotree: Using Online Social Networking to Strengthen Family Communication 359
Tsai-Hsuan Tsai, Yi-Lun Ho, Hsien-Tsung Chang, and Yu-Wen Li

HCI for Business and Innovation

It Was Nice with the Brick So Now I'll Click: The Effects of Offline and Online Experience, Perceived Benefits, and Trust on Dutch Consumers' Online Repeat Purchase Intention 371
Ardion Beldad and Mariël Segers

Interaction between Enterprise Resource Planning Systems and Organizational Culture 381
Carlotta Herberhold

Reconfiguring the Corporate and Commons: Mobile and Online Activism as a Form of Socio-technical Design 388
Constance Kampf

Social Media in the Product Development Process of the Automotive Industry: A New Approach 396
Andreas Klein and Götz Spiegel

Do All People Enjoy the Benefits from Technology Innovation? 402
Yoon Jeon Koh, Jae Heon Park, and Byung Do Chung

Semantically Integrated Business Applications for Enterprise Resource Planning Systems 411
Patricia Kraft and Rainer Thome

Towards a Next Generation Universally Accessible 'Online Shopping-for-Apparel' System 418
Kasper Kristensen, Nanna Borum, Line G. Christensen, Henrik W. Jepsen, Jacob Lam, Anthony L. Brooks, and Eva Petersson Brooks

Digital Menu Boards as Influencer for Healthy Eating 428
Anicia Peters and Brian Mennecke

Reverse Business Innovations – The Impact of ERP Software Upgrades on Organizations 438
Stefanie Rauff and Andreas Hufgard

Repository-Based Implementation of Information Pyramid: A Study Based on an ERP Case Study 446
Hans-Jürgen Scheruhn, Daniel Ackermann, Roman Braun, and Ulrich Förster

Agility Based on Stakeholder Interaction – Blending Organizational Learning with Interactive BPM 456
Christian Stary, Werner Schmidt, and Albert Fleischmann

Analysis of Customer Preference through Unforced Natural Passive Observation 466
Terumasa Tajima, Yusuke Iida, and Toshikazu Kato

The Development of an Innovative Design Process for Eco-efficient Green Products 475
Shiaw-Tsyr Uang and Cheng-Li Liu

Principled Ways of Finding, Analysing and Planning for Communicative Overhead in Interaction Technology for Fashion Industry 484
Jason Yang, Sean Rintel, and Stephen Viller

Continuing On-Premise or Adopt On-Demand? An Empirical Study of ERP Adoption in SMEs 492
Fan Zhao and Elias Kirche

Erratum

Exploring Children's Attitudes towards Static and Moving Humanoid Robots E1
Fang-Wu Tung and Tsen-Yao Chang

Author Index 501

Part I

Identity, Privacy and Trust

SP-CIP: A Secure and Privacy Aware Patient Centric Information Portal

Subrata Acharya

Department of Computer and Information Sciences, Towson University,
Maryland, United States
sacharya@towson.edu

Abstract. There is an urgent need for the establishment of patient-centric and quality aware solutions in today's healthcare environment. To this effect, Clinical Information Systems (CIP) play a powerful role to significantly influence the quality and efficiency of health care delivery systems. However, most of the current functional CIP are static [1] in design and typically represent and cater to only a single healthcare entity. This leads to an inaccurate and stale view of the patient status both in the care of the patient and also in the delivery of essential healthcare services. Moreover, it also delays patient care and leads to inconveniences in patient billing and the overall usability of the system. Added to this there is an oversight of security and privacy policies and procedures in the implementation of such systems. Thus, with the current pervasive nature of technology in such healthcare organizations, this creates a critical risk of such information systems to fall prey to data leaks and/or abuse during the access, storage and transmission of such systems.

To address the above key barriers towards the development of the effective CIP this research proposes *SP-CIP*: a standardized, security and privacy aware patient portal system. By utilizing the proposed standard design framework, the features of patient portals can evolve from static to a more dynamic design, which is inclusive of a user-centric, secure and proactive approach for patient care. These features also provide flexibility to include a de-centralized mechanism to access, store and transmit healthcare information over cloud based systems, and aims to establish the expectations of an effective patient portal, leading to improved cost and quality of patient care in today's healthcare systems. The pilot evaluation of the proposed systems at a high-end federal government Healthcare Information and Management Systems Society (HIMSS) 6 health care organization (> 500 bed capacity) and its associated partner organizations and clinical practices further demonstrates the strengths of proposed system.

Keywords: Clinical Information Portal, Security, Privacy Aware, Dynamic.

1 Introduction

With the dramatic increase in healthcare costs over the past few years many initiatives are being explored to address the associated challenges of providing acceptable quality with affordable cost of patient care. The current national administration's mandate

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 3–9, 2013.

to support, promote and foster electronic healthcare management systems is one such initiative. Furthermore, according to the Health Information Technology for Economic and Clinical Health (HITECH) Act, enacted as part of the American Recovery and Reinvestment Act of 2009, the goal is to promote the adoption and meaningful use of health information technology across the healthcare industry [2]. For instance, an estimated 30 billion dollars has been allocated for the use of electronic healthcare records (EHRs) in the Health Information Exchange (HIE) across the country [3].

Additionally, there are federal mandated guidelines and incentives for individual and group healthcare organization to adapt to electronic medium and to promote the meaningful use of their healthcare systems [4]. In this regard, healthcare information management and security is vital and imperative to maintain high quality of patient care, maintain reasonable costs as well as provide data confidentiality and address and mitigate the breach of personal data. To this effect, Clinical Information Portals (CIP) could play a very critical role to provide efficient and affordable quality healthcare to patients across the spectrum [5].

In this paper, we will first discuss the background of CIP and their current limitations *with regards to* today's technology. In section 3 will discuss the assessment of healthcare practitioners at a real-world Healthcare Information and Management Systems Society (HIMSS) 6 healthcare organization (> 500 beds) and 12 partner healthcare practices. To address the challenges we present the proposed SP-CIP Model in Section 4. In section 5 we discuss the pilot implementation of the proposed model and finally, conclude the paper.

2 Background and Current Limitations

A true clinical patient portal allows healthcare providers, patients and also insurance providers' access to electronic health information of the patient, typically through a web interface. In its current shape, patients can access their appointment schedules, current medications, test results, etc. Healthcare providers, can access patient health information to provide in-time and efficient care. This also allows out of network care providers to help their peers in collaborating and providing the necessary care to the patient as and when needed. CIP also allow patients to update contact information, provide their annual history information and send emails to clinicians and prepare them for a visit in a timely manner. It is also helpful for patient families to keep track of patient progress and monitor the health improvement of their near and dear ones. The insurance providers also connect to the CIP to provide quick decisions and timely billing to facilitate patient care and provide better coordination of care between healthcare providers and patients. In its current implementations the CIP do not address their design goals and hence it leads to all parties (healthcare providers, patients and insurance providers) to fall back on manual mechanisms of care, which defeats the purpose of establishment of CIP.

The current design of CIP is static and limited in their functionalities. This makes the use of such systems obsolete and creates delay in the overall care of the patient. We are also alarmed that there exists no consistent definition of the term portal itself

amongst the healthcare industry as a whole [6]. This creates additional interoperability issues when patient care is across distributed Health Information Exchanges (a typical scenario).

As part of the survey of both publically available and commercial CIP we found that there is no specific mapping of a particular type of information portal to the implementation environment. Also, with the surge of various technological means (such as desktops, laptops, PDAs, mobile phones, etc.) of access of the CIP there are severe constraints to the usability of such systems along with the issues of cost and security leading to vast discrepancies in the use of CIP to provide quality and cost effective patient care. Typically, CIP could operate in three modes, a) Server, b) Cloud and or c) Hybrid. Also, the critical security and privacy issues of the current portal design has led to non-use of these systems (more than 95% of the time), and users reverting back to manual operation.

Moreover, since the effective use of patient portals is over the current Internet, and since it operates on sensitive healthcare data; there are instance of significant cyber-attacks to such systems, and hence require substantial efforts to ensure they are secured appropriately. Our efforts in designing SP-CIP is to address the above said challenges and provide a flexible, secure and privacy aware model for the overall goal of effective healthcare care management to reduce cost and improve quality of patient care.

3 Assessment of Healthcare Practitioners and Target Patients

The goal of the assessment was to conduct an organization wide survey of healthcare practitioners and patients to understand the current limitations of CIP. We conducted our anonymous assessment on a selected uniformly distributed group of 50 healthcare providers (total 300 daily operational providers for the organization) and 30 in hospital care patients for the HIMMS 6 healthcare organization (> 500 beds).

Our goal was to evaluate the CIP design from the healthcare provider perspective based on the type of CIP implemented (sever, cloud based or hybrid), care location (in-hospital or out-patient care), interoperability, internetworking with other healthcare organizations, security and privacy of healthcare information. We also had mechanisms in place to measure per patient average time for care in the current target system. We realize that this is just the initiation of the assessment process and we will have other metrics to evaluate as we conduct periodic assessment and depending on the changing operational dynamics of the healthcare organization. Based on the feedback from the providers, we summarize our assessment as follows:

- The current server based design did not support the use of mobile technological devices used by the healthcare providers (physicians, physician assistants, nurses, lab technicians, etc). Due to the usability and security risks, the providers preferred and conducted manual (telephonic, in person) operation and disregarded the CIP.
- We noted that the attitudes of physicians favor server due to perception that their data is more “secure” on a server. There were also concerns of resource limitation in server base designs.

- Cloud base designs had the issue of HIPAA compliance and maintenance of the cloud, though the providers realized the advantages of real time access, flexible, low cost and resource availability over the cloud.
- The current design was not interoperable leading to healthcare providers reverting back to manual means to provide patient care.
- The associated clinical practices and other partner healthcare organizations did not connect in real time with the CIP of the central hospital causing delays in patient care for patient transfers, specialist consults, etc.
- Security and privacy metrics were not part of the CIP design. Also, there was lack of security assessment in the use and access of the CIP. This was the primary reason of lack of trust and non-usability of the system by a majority of the healthcare providers.

Next, our primary goal in assessing the patient use of the CIP was to understand the usability of the system and evaluate if the system aids their care by improving by addressing time and cost factors. Based on the feedback from the patients, we summarize our assessment as follows:

- Ignorance of the presence of such a system was the primary finding of the assessment.
- The usability and the system and the interactive nature were inefficient to cater to the patient care requirements, leading to patients reverting to walk-ins and telephonic interactions for all practical purposes.

We did conduct one-to-one discussions with a few insurance providers and collected statistics which also supported the above inferences on the current design of the CIP, which in effect rendered them obsolete.

To address all the above limitations, we have proposed the design of a secure and privacy aware portal system, SP-CIP model which we discuss in the next section. We successfully implemented the model in the pilot healthcare organization and it is currently in successful operation in the system.

4 SP-CIP Model

To address the above critical barriers towards the development of the effective CIP, this paper presents a standardized, secure and privacy aware patient portal system, SP-CIP. By utilizing the proposed design framework, the features of patient portals can evolve from static to a more dynamic design, which fits better in a real-time dynamic situation and leads to effective care of the patient. Additionally, it also incorporates a user-centric design to cater to user specific needs and requirements in the design of such CIP. Furthermore, the framework also incorporates secure design and privacy aware policies and procedures as part of the foundation and develops a more proactive approach towards patient care. Along with these features, the proposed framework also provides flexibility to include a de-centralized (distributed) mechanism to access, store and transmit healthcare information over cloud based systems, and aims to establish the expectations of an effective patient portal, leading to reduced cost and improved quality of patient care in future healthcare systems.

The SP-CIP model is based on a hybrid design (both server and/or cloud based). We have incorporated the technologies and secure protocols used in current secure web and wired and wireless communication as part of the design. A typical CIP node is represented as in Figure [1]. The salient functionalities are summarized as follows:

- Implementation of a secure multi-tier architecture that incorporates secure protocols, security measures, policies and procedures. The fundamental property of CIA (Confidentiality, Integrity and Availability) is verified for every information flow in and out for the system.
- The design isolates the web, application, and healthcare information servers behind multiple layers of secure Intrusion Detection and Prevention systems (IDS/IPS). There is also real-time monitoring and mitigation of developing vulnerabilities in the system. Our plan also includes periodic stress test to test the security and availability of the model.
- There is also secure mechanism of provisioning patient accounts on the CIP. SP-CIP also provides proper authentication, access control, and ensures the secure storage of the passwords. Standards for ensuring resets and adhering to providing security of *Least Privilege* in the design and access of the system.
- Implementing a proactive incident recognition and response program. Proactively monitoring for suspicious events, service interruptions, code errors, and general utilization issues.
- Hierarchical tiered design to ensure secure information flow between central hospital, peer healthcare practices, patient client devices and insurance providers systems. Adherence to current secure protocol standards and proactive mechanism to detect and mitigate potential anomalous conditions is mandated in the current design of SP-CIP.
- Adequate mechanisms for accountability via logging, log analysis and mitigation.
- Appropriate redundancy and recovery operations in place to ensure worst care operation and real-time availability of the system.
- Federal mandated compliance via periodic audits to ensure the appropriate operation of the proposed SP-CIP model.

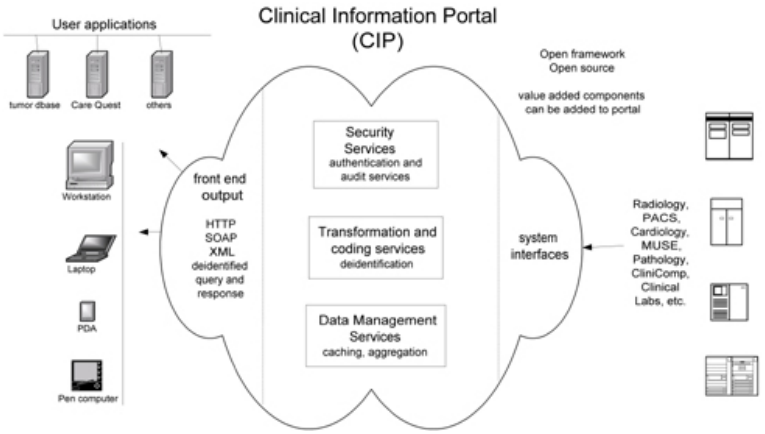


Fig. 1. SP-CIP Node

In the next section we will discuss the integration and pilot of the proposed SP-CIP model at the target HIMMS 6 healthcare organization. We also present the overall mechanism to incorporate the design of secure and privacy aware CIP as part of the standardization efforts to ensure security and privacy aware information flow in the overall healthcare organization.

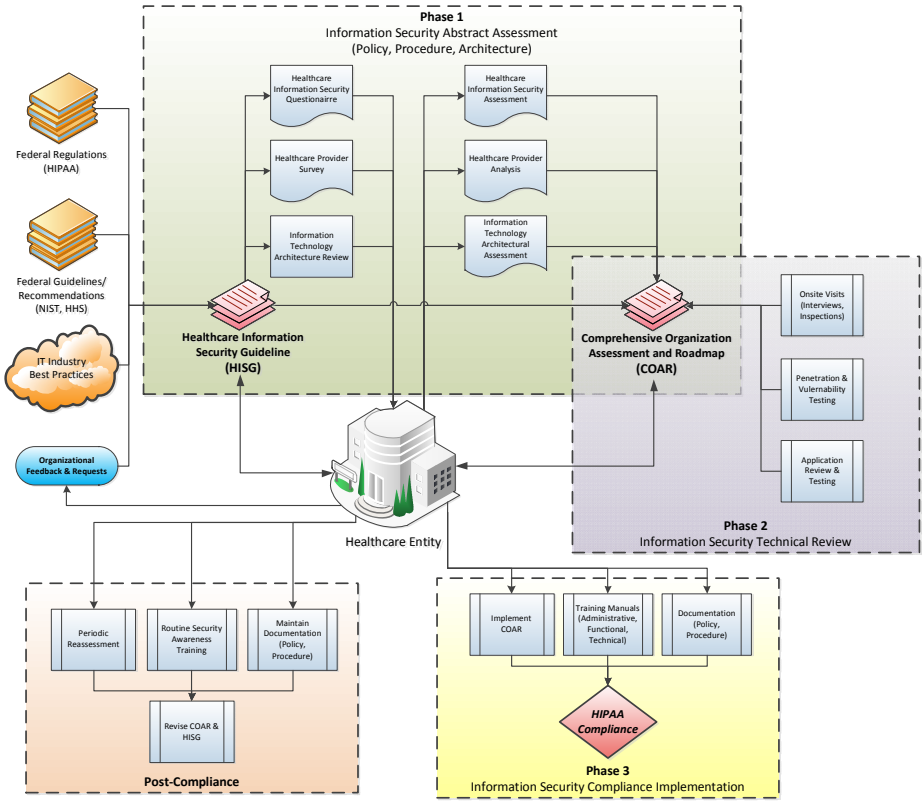


Fig. 2. Standardized framework for SP-CIP inclusion

5 Pilot Evaluation

Our proposed SP-CIP model has been currently piloted at a state-of-art federal government HIMSS 6 health care organization (> 500 bed capacity) and its associated 12 partner organizations and clinical practices. Figure 2 demonstrate the overall standardization framework which incorporated the SP-CIP design for the healthcare organization. The phases include the *initial assessment and evaluation*, the *technical review and implementation* and the *final compliance evaluation* and *periodic monitoring and enhancements*.

The overwhelmingly positive feedback from both the in-hospital patients and healthcare providers further demonstrates the strengths of the proposed system. Some of the measures for patient satisfaction were based on time to care, response to care, follow-ups on care and reminders, etc. The healthcare provider measures were based on overall improvement in patient average, ease of care and effective response to care situations, etc. Currently, we are in the process of conducting a comprehensive evaluation of the proposed model over the distributed network of the pilot healthcare organization and its partner institutions to demonstrate the improvement in cost reduction and effective patient care services. We also plan to design a checklist for laying out the "*best practices*" for effective CIP design, implementation and operation in collaboration with regulatory bodies at the federal government and share the proposed framework with other healthcare organization in and around the nation.

6 Conclusion

In this paper we propose SP-CIP as a vehicle to provide cost effective and quality aware patient care. Our model addresses all the current limitations and challenges by incorporating the federally mandated polices and the secure and privacy aware protocols and practices over the current WWW usage. We have been successful in providing improvements in cost and quality in our pilot organization and associated 12 partners. As future research directions we plan to incorporate the model in various national distributed healthcare networks and demonstrate the strength of the model.

References

1. iHealthBeat Patient Portal: Not Just another Static Text Field Form (2013), http://www.ihealthbeat.org/perspectives/2013/patient-portal-not-just-another-static-textfield-form.aspx (last accessed February 2013)
2. United States. Health Information Technology (HealthIT). Meaningful Use of EHR (2013), http://www.healthit.gov/policy-researchers-implementers/meaningful-use (last accessed February 2013)
3. Healthcare Information Security. $30 Billion: Money well spent (2013), http://www.healthcareinfosecurity.com/blogs/30-billion-money-well-spent-p-1185 (last accessed February 2013)
4. United States. Center of Medicare and Medicaid Services (CMS). EHR Incentive Program (2013), http://www.cms.gov/Regulations-andGuidance/Legislation/EHRIncentivePrograms/Meaningful_Use.html (last accessed February 2013)
5. Patient Portal. Patient Portals (2013), http://en.wikipedia.org/wiki/Patient_portal (last accessed February 2013)
6. Acharya, S., McSweeney-Feld, M.H.: Clinical Information Portals: Classification of Terminology and Design Implications. In: Academy of Business Research Fall 2012 International Conference, Atlantic City (September 2012)

Identity Management through "Profiles": Prototyping an Online Information Segregation Service

Julio Angulo and Erik Wästlund

Karlstad University,
Universitetsgatan 2, 651 88 Karlstad, Sweden
{julio.angulo,erik.wastlund}@kau.se

Abstract. Whereas in real everyday life individuals have an intuitive approach at deciding which information to disseminate to others, in the digital world it becomes difficult to keep control over the information that is distributed to different online services. In this paper we present the design of a user interface for a system that can help users decide which pieces of information to distribute to which type of service providers by allowing them to segregate their information attributes into various personalized profiles. Iterative usability evaluations showed that users understand and appreciate the possibility to segregate information, and revealed possible improvements, implications and limitations of such an interface.

Keywords: Usability, identity management, privacy preferences, partial identities, audience segregation, digital transactions.

1 Introduction

In everyday life individuals are frequently and naturally playing different social roles, as family members, professionals, hobbyist, activist, etc. Typically, people do not reveal all of their personal information to all of their different social groups, but rather they inadvertently and intuitively select the information that is suitable to share with members of a certain group depending on the context of a situation. Such phenomenon was first referred in the 1950's as *audience segregation* by sociologist Erving Goffman [15].

However, it can be claimed that in their 'digital lives' Internet users do not yet have the appropriate tools that help them manage their partial digital identities which let them segregate the information they distribute to different Internet services. For this reason, users tend to present similar identifiable attributes to many different service providers and to send more attributes than are actually needed to complete a transaction. Besides, nowadays users create different accounts with different services, which makes is hard for them to maintain and control which information is shared with whom. All these traces that users leave on the Internet could lead to higher probabilities of user impersonation, identity theft, profiling, and linkability attacks. A method for empowering users with control over their own personal information is needed as a way to minimize these risks.

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 10–19, 2013.

Therefore, in our work we try to tackle the challenge of designing an interface for what we refer to as *information segregation*, or the act of encapsulating different pieces of personal identifiable information in order to present only those attributes to deliberately chosen online service providers.

In this paper we first look at some work related to audience segregation and the usability of Identity Management (IdM) systems. We then present an overview of the process of design for a system supporting the idea of information segregation along with the performed usability evaluations. Finally we list some implications and conclusions about our work.

2 Related Work

Attention has been given to the challenge of letting users select the audiences to with whom they wish to share the content in social network sites (SNS), seen for instance in [6] [7] [12] [16] [18] [20] [29]. Google+ is a SNS that makes its audience segregation features explicit through the use of so called *circles*. It has been argued that Google+'s users have a "clear understanding of circles, using them to target information to those most interested in it" [31], and it can be said that the circles' interface offers further desired properties of interactive systems, such as consistency, playability, pliability, learnability, affordance, and others, as suggested, for instance, by Löwgren & Stolterman [21] and other design heuristics.

However, there is a distinction between the act of segregating audiences in SNSs, which creates a tension between the "desire for controlling our own information and the desire for unplanned social interaction" [17], and the act of distributing personal information for receiving commercial online services or products, which is motivated by a need or desire of the service or product being requested or the experience it provides.

For this reason, some attempts have been done at allowing users to act under different identities while communicating online. These include the efforts from Mozilla's Persona[1] and Google's Multiple Chrome Users[2], as well as the obsolete Microsoft's CardSpace and Firefox's plugin "Sxipper". Nevertheless, these existing systems are either at a proof-of-concept stage, have limited functionality, acting mainly as role-based access controlled or password management systems, or are hard to understand and therefore hard to adopt by regular users.

Regarding the usability of IdM systems, the research done by Jøsang et al. [19] discusses different models for IdM and suggest a user-centric approach for the management of user identities. Similarly, Eap et al. [13] recognize the need to provide users with more control over their identities distributed over different service providers, stating that IdM systems should assist users in their adoption of identity management practices. Moreover, a prototype called DRIM (Dresden Identity Management) [10] tried to integrate identity management concepts on Internet browsers, and subsequent work carried out as part of the PRIME

[1] `http://www.mozilla.org/en-US/persona/`
[2] `http://www.chromium.org/user-experience/multi-profiles`

project[3] has also realized the need for interfaces that support the notion of partial digital identities [23]. Furthermore, Dhamija & Dusseault [26] list seven concrete observations taken from their experience dealing with the design and analysis of security systems, trying to inform the reader why such systems often fail and ways to improve them. Alpár et al. [1] present some of the security, privacy and usability issues encountered in current IdM systems and propose recommendations for their improvement in each of these areas.

3 Conceiving an Interface for Information Segregation

3.1 Assumptions and Requirements

While conceiving an interface of a system that supported the idea of information segregation, we made assumptions about the architecture and the role that different parties would play in this envisioned system. For one, we assumed that an architecture is in place similar to the one suggested in the European projects PRIME and PrimeLife [8]. This implies that the system could be equipped with anonymous credentials technology (such as the one suggested by IBM's IdeMix [9]), and with a privacy policy matching engine (similar to the Platform for Privacy Preferences (P3P) [27] or the PrimeLife Policy Language (PPL) [5]) which can empower users to provide informed consent at the moment of releasing personal information.

However, in contrast to the architecture proposed in PrimeLife, it is assumed that the users' personal information is not stored locally on their devices, but instead it is stored centrally in a secure and privacy-friendly manner, for example in a cloud service provider.

It is also assumed that the proposed system can come in contact with different identity providers that are able to provide users with certified proofs of their identity. Similarly, online retailers would be able to offer their customers the possibility to complete a digital transaction using the certified or non-certified attributes contained within the customers' partial identities.

3.2 Design Approach

By following an iterative process of design, we created a series of sketches and lo-fi prototypes during three initial iteration cycles. The aim was to evaluate the use of different segregating metaphors, terminologies used, users' understanding of the concept of information segregation, as well as the intuitiveness of the interactive elements, look-and-feel of different design layouts and other visual aspects of the interface. Early evaluations actually revealed that participants consistently used the word *profiles* to refer to partial identities; therefore, we decided to give the name of "Profiles" to the IdM service being prototyped.

[3] PRIME - Privacy and Identity Management for Europe. `www.prime-project.eu`

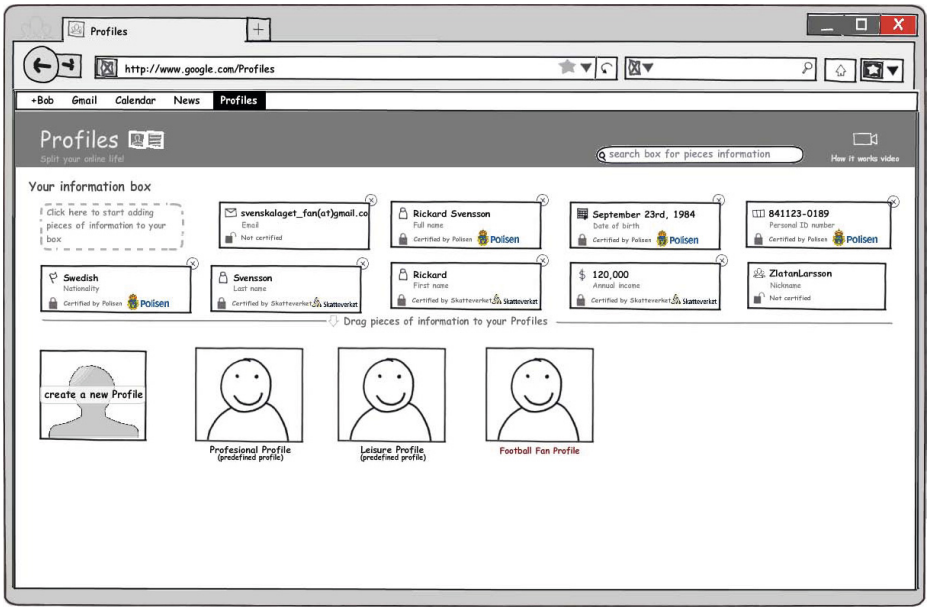


Fig. 1. A view of the "Profiles" prototype allowing users to manage their pieces of information at the top and their profiles at the bottom

Final Design Proposal. The final design proposal of the interface (depicted in Figure 1) can be briefly described as being divided into two panels. The top panel has a prominent label that reads '*Your information box*" conveying to users that it is a *place* where they can add, remove and manipulate their different pieces of information. Figure 2 represents an example of the look-and-feel of a piece of information, where users can see the attribute value of the information (the actual credit card number), the type of information (credit card), whether it is certified or not, and by which third party it has been certified. When users want to add a piece of information a dialog pops-up, where users can choose to either add non-certified attributes by typing them manually (Figure 3(a)) or to download certified attributes from a trusted identity provider (Figure 3(b)). As mentioned earlier, it is assumed that a list of trusted identity providers is already populated.

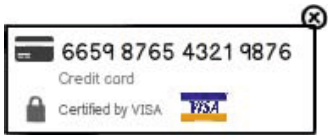
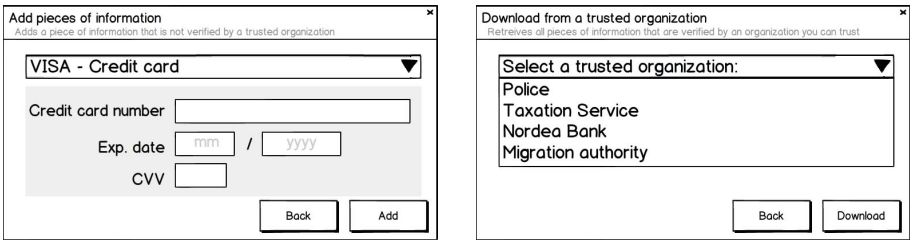


Fig. 2. The look of a piece of information

The bottom panel of the interface enables users to create and manage their different partial identities, or profiles. Results from the evaluations of earlier iterations indicated that users had a hard time giving relevant names to their profiles, therefore the interface provides a set of predefined profiles as a way to help first time users to get started and understand the purpose of the system. Profiles can be created and populated by clicking on the "create a new profile" icon, or by dragging one or many pieces of information onto them.



(a) Adding non-certified attributes manually

(b) Downloading certified attributes from a trusted identity provider

Fig. 3. Dialogs allowing users to add pieces of information

Every time pieces of information are added, users are given the option of selecting their preferences for the purposes for which service providers can use that information. For instance, users can specify that their credit card information should only be used for payment purposes. In this way, users can be informed if service providers respect their wishes during a digital transaction as explained in [3] [11]. Previous studies have shown the difficulty of users when stating their privacy preferences [11] [14] and to appropriately set privacy settings that would match their expectations to the reality of the protection of their privacy [22]. Given these observations, we realized through the different design iterations that the burden of setting such preference should be removed from users whenever possible, providing them with privacy-friendly default values as recommended, for instance, by data protection authorities. More importantly, the interface should promote the act of setting privacy preferences only in a moment that is relevant to the action at hand (as also suggested in [2]), for instance when information is added to a profile or when it is going to be distributed to service providers.

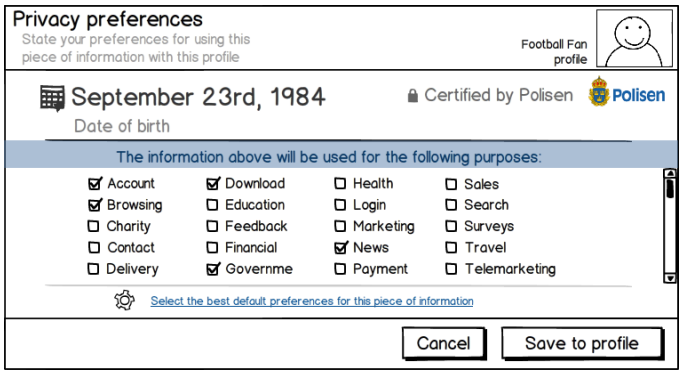


Fig. 4. Dialog allowing users to set the preferences of usage for the information pieces they add to a profile

3.3 Usability Evaluations

The initial design iterations were evaluated using different methods, such as eye-tracking technology, cognitive walkthroughs, thinking aloud protocols and questionnaires. Also, expert evaluations and feedback from professionals in the fields of e-commerce and identity management were taken into account. Results from these evaluations allowed us to identify possible improvements to the interface, and indicated that participants understood and appreciated the purpose behind information segregation.

For the final design iteration two main evaluation activities were carried out: usability tests using an interactive prototype and a focus group session. During the usability tests, all participants ($n = 12$) first answered a pre-questionnaire about their Internet habits and familiarity with existing audience segregation features in popular SNSs. Then, in the form of a cognitive walkthrough [24] they were asked to perform a series of tasks using an interactive prototype representing a scenario in which they were supposed to setup the "Profiles" system in order to make a purchase with an online service provider. The participants opinions and interactions with the prototype were noted as they went through each task, and they were also asked whether they considered a task to be easy or difficult to accomplish. At the end, participants answered a post-questionnaire including questions about the usability of the program, their understanding of its information segregation features, and PET-USES Likert-scale statements which aim to measure the secondary goals of users when interacting with privacy enhancing technologies, as presented in [30].

During the focus group session participants ($n = 30$) were first shown a demonstration of the "Profiles" interface and its privacy features. In order to encourage discussions and minimize 'group think' [28] participants were divided into smaller groups of 5 to 8 people. They were asked to discuss possible uses and improvements of the proposed interface with the other group members.

Summary of Results. From 12 participants that completed the cognitive walktrough phase, 8 had a very good intuitive idea of the purpose of the prototype, stating for example that "*the top lets you write pieces of your information, the bottom lets you create a profile with some of the pieces you wrote already in the top.*" Adding non-certified pieces of information was understood easier than downloading certified information, probably because users' unfamiliarity with the concept of certificates. However, after having downloaded certified information the first time, 11 out of 12 participants understood it well the second time they were asked to add certified information into the system. Observations also revealed that users were able to set privacy preferences easily, however in the post-questionnaire they reported this task as being the hardest to accomplish, suggesting a paradoxical view of privacy settings.

Participants from the focus groups session expressed, among other things, their concerns about centralizing all pieces of information in a single program: "*if a hacker gets into the program, then they will have all that information from me given by the Police, the Bank, etc.*". Moreover, they discussed the usage of

such a system that would fit their current situation and address their needs at hand. For instance, they envisioned a scenario where having a series of profiles could facilitate the way a user can look for different job positions that might require different information.

The complete list of tasks that participants were asked to go through, the instructions given during the focus group session and a detailed account of the results of the evaluations can be found in the technical report presented in [4].

3.4 Implications

Based on the process of design and the results from the evaluations, the following points present some of the major considerations to be taken when designing an IdM system like the one suggested hereby.

Progressive and Contextual Formation of Partial Identities. Since the early stages of design it became obvious that one of the major challenges would be to conceive an interface that would encourage users to start interacting with such a technology and adopt it for their routinary digital transactions. Thus, it is important to keep in mind that the users' first interactions with the system should be effortless, intuitive and perceived as useful at a moment that is relevant to what they are trying to achieve (e.g., purchasing a product through an online service). As stated in [26], "identity management is not a goal in itself", thus forcing users to, for instance, populate the system with many personal attributes so that it can work effectively during a purchase, will most likely discourage users to continue using it.

Instead, a progressive approach should be adopted in which users start by forming a partial identity in a simple way, gradually adding pieces of information as they are needed and creating identities at their own pace. The system should resemble to some extent the way identities are formed in real life and the contexts in which real identities evolve [1], consisting of progressive steps and determined by relevant life events, such as getting married, changing jobs, or opening a new bank account.

Maintaining Different Profiles. Creating different profiles could have the disadvantage that those profiles have to be maintained over time, which can become a burdensome task if users need to continuously monitor and update the information and the online services that they have associated with each profile.

A centralized solution would allow users to maintain their profiles across different devices over the air, which could provide convenience and further promote the gradual development of these partial identities. In this way, users would be able to access their profiles whenever and wherever they wish to do so, continuously evolving their identities 'on the fly' under relevant contexts of use in a progressive manner (a similar approach has been suggested during the PRIME and PrimeLife projects as presented in [3]).

Setting Privacy Preferences. Designing interfaces for letting users set their privacy preferences proved to be a difficult challenge. It was not until the final design suggestion that we thought we had come up with a simple enough mechanism to achieve this (seen in Figure 4). However, results from the post-questionnaire indicated that users still perceived this as a difficult step to complete.

What we have learned is that a system handling privacy preferences should relieve users from having to set those preferences from scratch, and instead a good set of privacy-friendly preferences, as defined by a trusted authority, should already be selected by default. The option to set or modify those preferences should not be made a priority, but should be made available within a context that a user will understand; for instance, at the moment of having to disclose information to a service provider, where she could, if she is interested or concerned, specify the purposes of usage of her data and other conditions.

4 Concluding Remarks

We have presented a design proposal for a system that enables users to group their information pieces into self-defined partial identities, or *profiles*. We refer to this act as information segregation. The suggested approach allows users to define the preferences for the data attributes contained within each profile, and to use those profiles at the moment of contacting specific applications or groups of online services with certain similarities, thus helping users protect their privacy by having a clearer approach to control their data and minimizing the personal information they disclose under certain application contexts.

We are aware of some of the limitations of the suggested interface. For one, difficulty of visualizing large amounts of pieces of information and managing more profiles than can be displayed on different screen screen sizes has not been fully considered. The level of complication can escalate when the formation of *sub-profiles* is taken into account (e.g., having a health profile that can be subdivided into pharmaceutical services, health clinic consultations, health clinic administration, etc.). Moreover, the users' understanding on how their information flows, where is it located (remotely or locally) and how it is handled by the IdM system is still unexplored (e.g., do users understand, or care about, what happens when they delete a piece of information?). Similarly, the possible steps to be taken to populate the system with trusted identity providers has not been considered. Additionally, we are aware of the privacy and security consequences that can arise from having all users' data in a centralized remote location; however, following a privacy-friendly architecture as presented in [25] can ensure users' privacy even towards the cloud service which stores the users' data.

Despite its limitations, we see this study as an exploratory approach towards useful privacy-friendly IdM systems. As part of future work we are working on creating scenarios where specific profiles might be employed (e.g., e-health or e-banking profiles) during a digital transaction. Also, we plan to work on adapting the interface to touch-screen devices of different sizes, and study the users' mental models of the location of their data as well as of the data flows

between devices and entities involved in a digital transaction. Moreover, in future iterations the mechanisms for specifying privacy preferences will be extended to not only include the purposes of data use, but also data retention periods, deletion obligations and other similar conditions.

Acknowledgments. This work is funded by a Google Research Award and the Swedish Knowledge Foundation (KK-stiftelsen) through the U-PrIM project. We would like to thank our colleagues Martin Ortlieb, Stephan Micklitz, Peter Gullberg, Simone Fischer-Hübner and Tobias Pulls.

References

1. Alpár, G., Hoepman, J.H., Siljee, J.: The identity crisis. Security, privacy and usability issues in identity management. Computer Research Repository (CoRR) (2011)
2. Angulo, J., Fischer-Hübner, S., Pulls, T., König, U.: HCI for Policy Display and Administration. In: PrimeLife - Privacy and Identity Management for Life in Europe, ch. 14, p. 261. Springer (June 2011)
3. Angulo, J., Fischer-Hübner, S., Pulls, T., Wästlund, E.: Towards usable privacy policy display & management. Information Management & Computer Security 20(1), 4–17 (2012)
4. Angulo, J., Wästlund, E.: Identity Management for online transactions - Using "Profiles" to segregate personal information. Tech. rep., Karlstad University, Karlstad, Sweden (April 2012), `http://www.is.kau.se/julioangulo/angulo/publications/2012/2012_IdentityManagementForOnlineTransactions.pdf`
5. Ardagna, C.A., Bussard, L., Di, S.D.C., Neven, G., Paraboschi, S., Pedrini, E., Preiss, S., Raggett, D., Samarati, P., Trabelsi, S., Verdicchio, M.: Primelife policy language. In: Proceedings of the W3C Workshop on Access Control Application Scenarios, Luxembourg (November 2009)
6. van den Berg, B., Pötzsch, S., Leenes, R., Borcea-Pfitzmann, K., Beato, F.: Privacy in social software. In: Camenisch, J., Fischer-Hübner, S., Rannenberg, K. (eds.) Privacy and Identity Management for Life, pp. 33–60. Springer, Heidelberg (2011)
7. van den Berg, B., Leenes, R.E.: Audience Segregation in Social Network Sites. In: Proceedings for the Second IEEE International Conference on Social Computing/Second IEEE International Conference on Privacy, Security, Risk and Trust, pp. 1111–1117. SocialCom/PASSAT, SSRN, Minneapolis, USA (2010)
8. Camenisch, J., Fischer-Hübner, S., Rannenberg, K.: PrimeLife - Privacy and Identity Management for Life in Europe, 1st edn., vol. 14. Springer (June 2011)
9. Camenisch, J., van Herreweghen, E.: Design and implementation of the idemix anonymous credential system. In: Proceedings of the 9th ACM Conference on Computer and Communications Security, pp. 21–30. ACM (2002)
10. Clauß, S., Kriegelstein, T.: Datenschutzfreundliches identitätsmanagement. Datenschutz und Datensicherheit 27(5) (2003), `http://dblp.uni-trier.de/db/journals/dud/dud27.html#ClaussK03`
11. Cranor, L.F., Guduru, P., Arjula, M.: User interfaces for privacy agents. ACM Trans. Comput.-Hum. Interact. 13(2), 135–178 (2006)
12. DiMicco, J.M., Millen, D.R.: Identity management: multiple presentations of self in facebook. In: Proceedings of the 2007 International ACM Conference on Supporting Group Work, GROUP 2007, pp. 383–386. ACM, Sanibel Island (2007)

13. Eap, T.M., Hatala, M., Gasevic, D.: Enabling user control with personal identity management. In: IEEE International Conference on Services Computing, SCC 2007, pp. 60–67. IEEE, Salt Lake City (2007)
14. Fischer-Hübner, S., Pettersson, J., Bergmann, M., Hansen, M., Pearson, S., Mont, M.: Human-Computer Interaction. In: Camenisch, J., Leenes, R., Sommer, D. (eds.) Digital Privacy. LNCS, vol. 6545, pp. 569–595. Springer, Heidelberg (2011)
15. Goffman, E.: The presentation of self in everyday life. Doubleday (1959)
16. Gonçalves, J.: Groupster: Narrowcasting on Social Networking Sites. Master's thesis, Madeira Interactive Technologies Institute, University of Madeira (2011)
17. Grimmelmann, J.: Saving Facebook. Iowa Law Review 94(4), 1137–1206 (2009)
18. Gross, R., Acquisti, A.: Information revelation and privacy in online social networks. In: Proceedings of the 2005 ACM Workshop on Privacy in the Electronic Society, WPES 2005, pp. 71–80. ACM, New York (2005)
19. Jøsang, A., Zomai, M.A., Suriadi, S.: Usability and privacy in identity management architectures. In: ACSW Frontiers, pp. 143–152 (2007)
20. Kairam, S., Brzozowski, M., Huffaker, D., Chi, E.: Talking in circles: selective sharing in google+. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2012, pp. 1065–1074. ACM, New York (2012)
21. Löwgren, J., Stolterman, E.: Thoughtful Interaction Design: A Design Perspective on Information Technology. MIT Press (2007)
22. Madejski, M., Johnson, M., Bellovin, S.M.: The failure of online social network privacy settings. Tech. rep., Columbia University (2011), http://www.futureofprivacy.org/wp-content/uploads/2011/07/TheFailureofOnlineSocialNetworkPrivacySettings.pdf
23. Pettersson, J.S., Fischer-Hübner, S., Danielsson, N., Nilsson, J., Bergmann, M., Clauss, S., Kriegelstein, T., Krasemann, H.: Making PRIME usable. In: Proceedings of the 2005 Symposium on Usable Privacy and Security, SOUPS 2005, pp. 53–64. ACM, New York (2005)
24. Polson, P.G., Lewis, C., Rieman, J., Wharton, C.: Cognitive walkthroughs: a method for theory-based evaluation of user interfaces. International Journal of Man-Machine Studies 36(5), 741–773 (1992)
25. Pulls, T.: Privacy-friendly cloud storage for the data track. In: Jøsang, A., Carlsson, B. (eds.) NordSec 2012. LNCS, vol. 7617, pp. 231–246. Springer, Heidelberg (2012)
26. Dhamija, R., Dusseault, L.: The Seven Flaws of Identity Management: Usablity and Security Challenges. IEEE Security & Privacy Magazine 6(2), 24–29 (2008)
27. Reagle, J., Cranor, L.F.: The platform for privacy preferences. Commun. ACM 42(2), 48–55 (1999)
28. Rubin, J., Chisnell, D.: Handbook of usability testing : how to plan, design, and conduct effective tests. Wiley Publ., Indianapolis, Ind. (2008)
29. Tootoonchian, A., Saroiu, S., Ganjali, Y., Wolman, A.: Lockr: better privacy for social networks. In: Liebeherr, J., Ventre, G., Biersack, E.W., Keshav, S. (eds.) CoNEXT, pp. 169–180. ACM (2009)
30. Wästlund, E., Wolkerstorfer, P., Köffel, C.: PET-USES: Privacy-enhancing technology – users' self-estimation scale. In: Bezzi, M., Duquenoy, P., Fischer-Hübner, S., Hansen, M., Zhang, G. (eds.) IFIP AICT 320. IFIP AICT, vol. 320, pp. 266–274. Springer, Heidelberg (2010)
31. Watson, J., Besmer, A., Lipford, H.R.: +your circles: sharing behavior on google+. In: Proceedings of the Symposium on Usable Privacy and Security, SOUPS 2012, pp. 12:1–12:9. ACM, New York (2012)

An Empirical Investigation of Similarity-Driven Trust Dynamics in a Social Network

Yugo Hayashi[1], Victor Kryssanov[2], and Hitoshi Ogawa[2]

[1] Faculty of Library, Information and Media Science, Tsukuba University
1-2, Kasuga, Tsukuba, Ibaraki 305-8550, Japan
hayashiyugo@gmail.com
[2] Faculty of Information Science and Engineering, Ritsumeikan University
1-1-1 Nojihigashi, Kusatsu 525-8577, Japan
{kvvictor,ogawa}@is.ritsumei.ac.jp

Abstract. Presently, people often create and keep lists of other people with similar preferences for hobbies, such as books, movies, music, and food in online social network service systems. Recent studies in recommender systems have shown that the user's data can be used to recommend items based on other users' preferences (e.g. as implemented in amazon.com). To make such systems more effective, there is a need to understand the mechanism of human trust formation. The goal of this study is to develop cognitive models describing the trust formation in social networks. This paper presents results of a controlled experiment conducted to collect human behavior data through a series of trust evaluation tasks.

Keywords: Social Cognition, Social networks, Trust Dynamics, Recommender Systems, Human-Computer Interaction, Conversational Agents.

1 Introduction

Studies in recommender systems have suggested that a user's data can be used to recommend this user items, based on other users' preferences. Although, a relevant question comes up: Can one predict how people may rely on and/or trust opinions of others in their own decision-making? To develop efficient and effective recommender systems, there is a need to understand the mechanism of human trust formation. The paper proposes a new experimental design where computer agents are used to capture the dynamics of trust formation through interactions in a social network.

1.1 Similarity and Trust in a Social Network

Studies in social psychology showed that trust is often related with similarity [1]. Studies in human and computer interaction, on the other hand, showed that users tend to prefer recommendations from friends rather than from computer systems [2]. An experiment conducted by [3] showed that the profile similarity may be related to the

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 20–28, 2013.

ways the users determine whether to trust other users when solving an on-line selection task. The study provided implications for how to predict trust, and how the corresponding model would be incorporated into related algorithms, such as collaborative filtering algorithms used in recommender systems. The study focused on the relationship between similarity and trust but the proposed model does not address the dynamics of user similarity in an online-social network.

Given this background, two relevant questions arise: 1) how could one develop a behavioral model of the similarity and trust development for the application in recommender systems, using information extended from social networks, and 2) can one simply use similarity as an index of trust? The presented study investigates these questions in an open-ended environment, where individuals interact through time.

1.2 Recommendations by Other Users in a Social Network

In social network studies, there were attempts to develop recommender systems and algorithms based on the users' similarity and trust (e.g. [4], [5], and [6]). A popular method to provide the user with adequate recommendations is the so-called collaborative filtering. Collaborative filtering allows for recommending items to users, based on their similarity and/or on the similarity of the items the users prefer to other items. A more advanced method for recommending items to the users is the Bayesian filtering algorithm. Using the latter algorithm, it is often possible to predict the dynamics of human trust, based on the past results of trust/distrust and similar/dissimilar evaluations. There were, however, few studies that would try to validate the Bayesian framework with real human data. As many studies in psychophysics showed, psychological states related to trust may differ due to differences in physical world. To address this problem, the presented study focuses on an experimental situation where an individual perceives a single dissimilar physical state while interacting with other users.

1.3 This Studies Objective

This study thus aims to explore how human participants react towards the recovery of trust throughout the interaction process. It investigates differences between a simulation model currently used in many recommender systems and the corresponding human performance observed in an experiment. The commonly used Naïve Bayes model is validated with empirical data as a correct model of the human trust dynamics. Data generated with the Bayes model is used as the criterion for comparison with empirical human data for understanding parameters of the trust development in a social network. The following three issues are examined in the study:

1. How similar/dissimilar preferences affect trust dynamics in an (online) social network?

2. Assuming that after some time, similar preferences may contribute to the development of a high trust and a relationship, then when one perceives a dissimilar state, what is the immediate dynamics of trust?
3. How would the empirical data differ with the corresponding normative behavioral data obtained through simulation with the classical Naïve Bayes model?

2 Method

A web-based system has been developed for conducting a controlled experiment to understand cognitive aspects of trust in a social network. In the proposed experimental paradigm, a multi-conversational agent was used as an imaginary actor playing the role of a member of the social network. This experimental method helped us understand specific stimulatory responses, as observed in trust dynamics, when an unexpected incidence occurs in the social network.

2.1 Procedures

In the experiment, participants (who are all members of one social network) evaluated opinions (preferences) of other participants for a topic provided. More specifically, they watched a video-clip of a short cartoon, using smart phones, and shared their impressions about the cartoon with other participants through a web-page. The procedure was as follows: 1) Participants first watch the video, using smart phones, 2) They evaluate how interesting they felt about it, 3) They see the other users' evaluation results and, finally, 4) They evaluate how strong they 'trust' the other users (see Fig. 1).

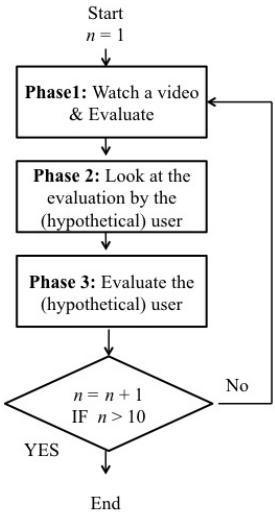


Fig. 1. The experimental procedure

The procedure was continued for about 10 trials (the exact number of trials depended on the experiment). During the task, the participants watched 10 different video clips in a predefined order, evaluated their impressions, and shared the evaluation results with each user in every trial. Results from users were manipulated by computer agents (i.e. by the hypothesized users), and were controlled to change over time. This methodology is adopted from a previous study [7]. One of the hypothesized users' evaluations was adjusted to be always almost the same as the participant's, excepting for one trial where the participant experienced a "shock" event due to completely different evaluations learned from the other users. Also, participants were asked to do a secondary evaluation by evaluating the other users, based on how strong they "trusted" the other users' opinions. Details of the experiment are discussed in the following sections.

Phase 1. First, participants accessed the Web-based application using their mobile devices. In beginning of the task, all participants received the following instruction: "You are randomly connected to four persons in this experiment, who are also now watching the cartoon and giving their evaluations." Actually, there were, however, no other real users but computer conversational agents responding appropriately to each participant. This instruction made participants to think that they interact with members of the social network in real time. The participant starts the task by rating on a collection of short films by accessing You Tube (`http://www.youtube.com`). Episodes of a short cartoon "Tom and Gerry" were used as the stimuli. After the participant finished watching one clip he or she evaluated it, based on three questions (see Fig. 2). Each evaluation was on a ten-point lickert-scale, and the three questions were: "how good was the content", "how good was the character", and "how good was the scenario" (as translated from Japanese).

Fig. 2. Screenshots from the ratings phase

Phase 2. After the participant finished rating, he or she can see how other users rated the episode. All other users' ratings were generated by computer agents. These latter ratings will be called hypothesized users' "rating profiles". Fig. 3 shows an example of the rating profiles.

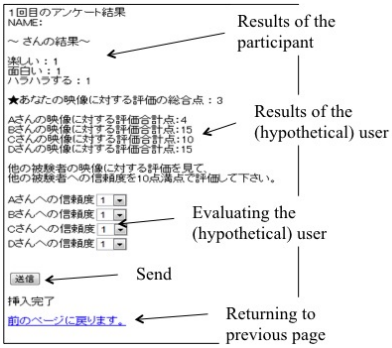


Fig. 3. Screenshots from the rating phase

Two types of the results were shown: (1) the sum of ratings by the participant and (2) the sum of the rating profiles by the hypothesized users. Let us denote the participant's sum of the ratings of the episode in trial i as Si, the hypothesized user's sum of rating profiles as U_a,U_b, U_c, U_d, and the absolute difference between any two ratings as β. For each episode i, β was randomly selected from the following three categories by the computer agents:

- Small variation : $0 \leq \beta \leq 2$,
- Medium variation: $3 \leq \beta \leq 5$,
- Large variation: $6 \leq \beta \leq 9$.

To investigate how similar/dissimilar preferences affect the trust dynamics, one of the four hypothesized members labeled "User A" was always adjusted to generate small variations. Two other hypothesized members (labeled "User B" and "User C") always generated large variations, and one such member generated medium variations ("User D"). If people trust more to the users with similar ratings, it would be natural to expect that a subject would develop higher trust towards "User A" than to the other users. It also appears natural to expect that this tendency would increase over time.

Table 1. Rating profiles of users and trials

Hypothetical user	Trial i	β
U_a	1, 2, 4, 6, 7, 8, 9, and 10	Small variations
	3, 5	Large variations
U_b	All trials	Large variations
U_c	All trials	Medium variations
U_d	All trials	Large variations

User A was manipulated almost in the same manner as the participant behaved, excepting for trial 3 and 5. In that trial, User A response, U_a, was adjusted to change from Small variations to Large variations. Through this adjustment, the participant experienced a shock, when the result of the evaluation changed dramatically compared to other trials. U_b and U_d were adjusted to change always in Large

variations, and U_c was adjusted to change always in Medium variations. The data of the secondary evaluation of trust was collected and used as the main data for the analysis.

Phase 3. In this phase, the participants were asked to rate how much they trusted the hypothetical users, based on what they observed in the rating profiles. Participants evaluated four hypothetical users on a ten-point lickert scale. After finishing this activity, they proceeded to the next trial, and watched a new video clip. This cycle continued until the participants finished 10 trials.

2.2 Participants and Condition

Two experiments were conducted to investigate the cognitive aspects of trust in a preference evaluation task. In both experiments, all the participants were students enrolled in a psychology class. The participants were asked to participate in the experiments for course credits. The experiments took place in the classroom and the participants were told to use their smartphones in the experiments. In the experiment, twenty-seven participants (male = 7, female= 20) made evaluations in ten trials.

3 Results

3.1 Overall Result

To investigate the trust dynamics for a longer period of time, Experiment 2 was conducted. Results obtained in this experiment were then analyzed, using a 4 (users) x 10 (trials) within-subjects factorial design. Fig. 4 shows the results of the trust evaluations for the hypothetical users. The vertical axis gives the average of the trust evaluations, and the horizontal axis represents the trials.

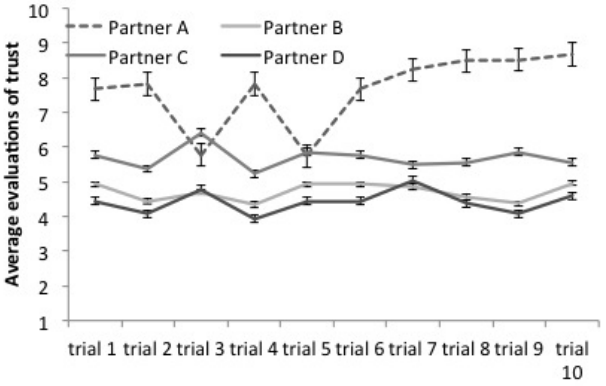


Fig. 4. Results of evaluations

The ANOVA analysis revealed that there was an interaction between the two factors ($F(26,702)=4.768$, $p < .01$). Simple main effects in the users detected several differences among trials. Differences were found in all trials ($F(3,936)=19.753$, $p < .01$; $F(3,936)=27.758$, $p < .01$; $F(3,936)=6.624$, $p < .01$; $F(3,936)=29.880$, $p < .01$; $F(3,936)=4.430$, $p < .01$; $F(3,936)=25.505$, $p < .01$; $F(3,936)=24.200$, $p < .01$; $F(3,936)=35.359$, $p < .01$; $F(3,936)=40.421$, $p < .01$; $F(3,936)=33.805$, $p < .01$). Multiple comparisons, using Ryan's method showed that User A was evaluated higher than users B, C and D in trial 1 ($p < .01$; $p < .01$; $p < .01$). These results were consistent for trials 1, 2, 4, 6, 7, 8, 9, and 10. The result of the two experiments has, therefore, demonstrated that the ratings of trust become higher when the participants observe a similarity to their own opinions in the shared opinions about the cartoon. This finding generally confirms that similar preferences enhance trust. The results obtained also showed that the ratings of trust dropped as the participants experienced the shock of non-similar evaluations. This confirms that similarity of preferences and trust are strongly correlated, and this correlation can be enhanced over time.

3.2 Simulation

A further analysis was conducted, using the Bayesian model for a comparison with the data. Evaluations towards User A were used for investigating how trust recovers after experiencing the shock. In the analysis, evaluations of User A were binary coded to either, "trust" (1) or "no trust" (0). After the second trial, when participants rated lower than in the previous trial, the evaluating were coded as "not trust". When participants rated higher or the same, compared to the previous trial, the evaluations were coded as "trust". The ratio of selecting "trust"/"not trust" was calculated for each trial. Using this coding scheme, all the data was then compared with the Bayesian model. For the analysis, prior probability of perceiving a similar /dissimilar evaluation in trial i was determined as H_i, and the probability of generating trust was $P(H_i|D)$. The following equation specifies the Naïve Bayesian model used:

$$P(H_i|D) = \frac{P(D|H_i)P(H_i)}{P(D|H_1)P(H_1)+P(D|H_2)P(H_2)} \quad (1)$$

$$where\ \mathrm{i} = 1, 2. \quad (2)$$

Fig. 5 shows the empirical data in comparison with simulation results. The vertical axis gives the probability of trust, and the horizontal axis represents the trials. The results demonstrate that the Naïve Bayes model fails to accurately predict the speed of recovering from distrust to trust but still qualitatively is in a good agreement with the observed empirical dynamics.

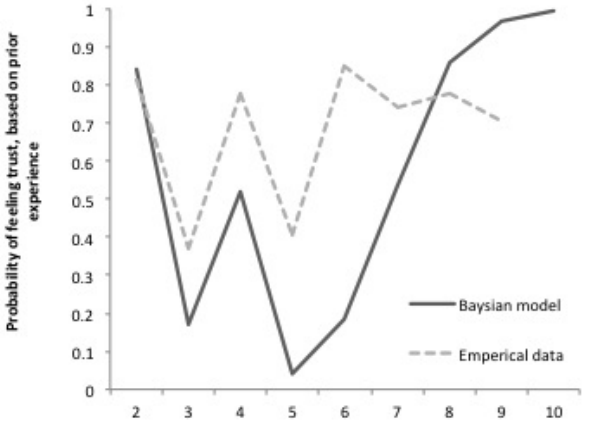


Fig. 5. Results of the simulation

4 Discussion and Conclusions

The trust dynamics observed in the experiments revealed that when one perceives a dissimilar state, the subjective perception of trust decreases temporary, but then almost immediately recovers upon a positive experience. This fact is interesting in the view of understanding the users of various social network service systems, who tend to ignore social conflicts and keep developing their trust in respect to a particular member, while ignoring negative experience of interactions with that member. This phenomenon would be attributed to so-called conformation bias. Several studies have recently shown that people are likely to become biased to (mis)trust others' opinions in an online environment (e.g. [8]). It follows from the results obtained in our study that such bias on trust may rapidly develop in social network service systems. This finding would have implications for design of recommender systems based on information extracted from social networks. In future work, we plan to investigate in detail the relationships connecting trust and similarity to find ways for controlling the conformation bias.

References

1. Byrne, D.: The attraction paradigm. Academic Press, New York (1971)
2. Sinha, R., Swearingen, K.: Comparing recommendations made by online systems and friends. In: Proceedings of the DELOS-NSF Workshop on Personalization and Recommender Systems in Digital Libraries. Dublin, Ireland (2001)
3. Golbeck, J.: Trust and Nuanced Profile Similarity in Online Social Networks. ACM Transactions on the Web 3(4-12), 1–32 (2009)
4. Walter, E.F., Battiston, S., Schweitzer, F.: A model of a trust-based recommendation system on a social network. Journal of Autonomous Agents and Multi-Agent Systems 16(1), 57–74 (2007)

5. Ziegler, C.N., Lausen, G.: Spreading activation models for trust propagation. In: Proceedings of the IEEE International Conference on e-Technology, e-Commerce, and e-Service, pp. 83–97. IEEE Computer Society Press, Los Alamitos (2004)
6. Ziegler, C.N., Golbeck, J.: Investigating Correlations of Trust and Interest Similarity. Decision Support Systems 42(2), 460–475 (2006)
7. Hayashi, Y.: The effect of "Maverick": A study of Group Dynamics on Breakthrough in Collaborative Problem solving. In: Proceedings of the 34th Annual Conference of the Cognitive Science Society, pp. 444–449. Lawrence Erlbaum Associates, Hillsdale (2012)
8. Joinson, A.N.: Understanding the psychology of internet behavior: virtual worlds, real lives. Palgrave, Basingstoke (2003)

The di.me User Interface: Concepts for Sharing Personal Information via Multiple Identities in a Decentralized Social Network

Fabian Hermann[1], Andreas Schuller[1], Simon Thiel[1],
Christian Knecht[1], and Simon Scerri[2]

[1] Fraunhofer-Institute for Industrial Engineering, Stuttgart, Germany
{firstname.lastname}@iao.fraunhofer.de
[2] Digital Enterprise Research Institute, National University of Ireland, Galway, Ireland
simon.scerri@deri.org

Abstract. The di.me userware visualizes vast personal information from various sources and allows for sharing them in a decentralized social network. Multiple identities can be used to avoid unintended linkability when communicating to other users or external systems. The di.me user interface for that is presented in this paper. A user-centered information- and view-structure for items of the personal information sphere and concepts for multiple identities used for sharing are discussed.

Keywords: user interface design, decentralized social networking, personal identity management, security and privacy.

1 Introduction

Personal information is widely used for social networking and collaboration in business and private life. The project digital.me[1] researches on intelligent management of personal information and its use for social networking. A developed framework and system prototype – the di.me userware – shall integrate the personal information sphere by crawling and synchronizing data from various sources, e.g. files from personal storages and devices, profiles or contacts from external social networks. In order to achieve a high user control on personal data, the system architecture is conceptualized as a decentralized social network (Scerri et al., 2011), which allows to run *personal nodes* of the di.me userware either as self-hosted server or as user account on a provided server. Compared to o other decentralized social networks (Thiel et al., 2012), di.me follows an extended approach by providing intelligent features like context-aware recommendations, trust advisory (Heupel et al, 2012., in press), or rule-based automation of user tasks (Scerri et al., in press).

The personal information in di.me is represented by a set of *ontologies* (Scerri et al., 2012). Ontological knowledge is used to power the di.me semantic core in

[1] www.di.me-project.eu

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 29–38, 2013.

providing the automatic integration of distributed personal information, its improved management, as well as the recommendations and user advisory.

A specific approach of di.me is to provide *multiple, partial identities* (Pfitzmann, Hansen, 2010), that can be used to communicate or share information with other di.me users or external systems. For the design multiple-identity-systems, it is an important guideline (Hansen et al. 2008) to avoid linkability: Partial identities shall not be linkable to the same sender if not intended. In single-identity-systems, linkability risks are high as they support just one communication identifier, even if pseudonymized. I users want to have distinct identities in single-identity-systems, they must fall back to multiple user-accounts in the system (which e.g. Skype suggests[2]). In contrast, multiple-identity-systems link partial identities to a root identity. To technically enable continuity in communication, communication identifiers have to be used. In di.me, these identifiers are the "service-accounts" (for a discussion of service-accounts and anonymity requirements on the network level see Bourimi et al., 2012). Identities and claims in di.me are unproven, i.e. di.me does not act as an identity-provider. However, the integration of proven identities with anonymous credential systems was demonstrated (Schwarte et al, in press).

The di.me user interface (UI) is realized as web UI and a mobile Android client. The major challenge for the UI conceptualization and design presented in this paper was to find a flexible model that enables interaction the personal information sphere and to use them in a decentralized social network. A central research question was how to interact securely with the partial identities, and which concepts or metaphors for that shall be used in the UI.

2 UI Concepts for Personal Information and Identity

As social systems are highly developed, for many aspects relevant UI design patterns or established design solutions can be applied to achieve a self-explanatory UI. A prevalent pattern used in social systems are *groups of contacts* (see e.g. Crumlish, Malone, 2009) which resemble closed networks of people and allow the user to describe the relationship with individual contacts. This pattern was prominently marketed as added-value feature by Google+[3] ("circles") in 2011 (Simonite, 2011). Another example are "aspects" in the Diaspora network[4]. Groups strongly interplay with personal identity management, as groups can be used to control disclosure of partial information. The concept "group" may be simply a category of contacts of an individual user only or can be shared with others (Schümmer, Lukosch, 2007), e.g. for user management on centralized platforms. The group concept in OpenSocial[5], e.g. is a shared o centrally managed group for discussion forums.

The identity shown to others, often are called the "profile" (Crumlish, Malone, 2009), being the complete identity (Pfitzmann, Hanssen, 2010), i.e. the superset of all

[2] https://support.skype.com/en/faq/fa829/
[3] plus.google.com
[4] diasporaproject.org
[5] opensocial.org

personal attributes which may be separated by privacy options. For *partial identities (sets of attributes)* that can be shown to others, different metaphors are used. In Microsoft CardSpace "cards" are shown which include attributes, potentially a certificate, and a recipient-specific, non-global identifier (Al-Sinani et al., 2010). The UI guides the user to select the same card for a recipient each time (Steuer et al., 2010). In the identity management of some providers of OpenID[6], attribute sets are presented as "personas". Users can manage personas, e.g. "work" with different attributes like e.g. nickname, or e-mail address. The term "persona" is also used by the Mozilla identity management [7] of web site credentials. "Identity" stands for attributes attached to the global identifier in the OpenID UI, probably indicating that the global identifier is the superordinate concept to "personas". A conflicting meaning of "identity" is used in a Mozilla Thunderbird extension[8]: Here, the attribute sets are called "identities", and they may be used within one email account. These and other examples may indicate that up to now, no common understanding and use of identity-related concepts has evolved within the UI-design for end-user applications.

Non-global communication identifiers are needed for continuity in communicating in multi-identity-systems. We consider them as less known to end-users, as many end-user-systems are single-identity-systems (even if offering partial sets of attribute). This may increase the difficulties of users with building up mental models of identity management systems (Sun et al, 2011). However, one case of multiple identifiers which is prevalent in end-user-systems are multiple email accounts. Email-clients like Microsoft Outlook or Mozilla Thunderbird offer selection of the account used for a message by a "From"-field in the UI.

3 The di.me UI for Personal Information and Identities

The di.me UI enables the user to interact with the personal information sphere and communicate via multiple identities. For that, a UI object model based on the semantic knowledge representations was developed and a view structure which focused on a web UI but also considered the adaption to the Android client.

They resulted from a *user-centered design process* with two major concept versions, each tested by a usability test: A first concept version was realized through a wireframe mockup without any visual design. It was tested in the first usability test. 21 participants from potential di.me user groups (private end-users and business users working with social system e.g. as public relations managers) worked in moderated sessions on predefined tasks. The method was a formative evaluation (Lewis, 2012). The qualitative results of observation, think aloud, and interviews were aggregated to categories considering the underlying causes of user problems (Dumas, 2003). Additional summative evaluation measures were used. Based on these results and further technical requirements, a second concept version was realized through a partly functional HTML-prototype with accurate visual design, and a functional Android client.

---

[6] openid.net

[7] mozilla.org/persona

[8] addons.mozilla.org/thunderbird/addon/flexible-identity

It was tested in the second usability test with 20 participants, mainly private end-users. The methodology used was similar to the first test.

In this section, we present the UI concepts for personal information and identity, and qualitative test results related to that.

3.1 UI Objects

The object model defines the scope of information and the terminology applied in the UI, and by that plays a key role for the perception and usability of the system by the end-user. A specific challenge was to find object definitions that are appropriate to convey not only the proprietary functionality but also the information synchronized from external services. Besides the core concepts presented below, additional concepts were e.g. "Message", referring to direct communication messages to a contact or group, "Service" referring to the external systems that are being connected to a di.me personal node, or "Situation" which refers to a high-level description of the user's context for the context-related features.

The UI objects interplay with the semantic knowledge representation: They can be mapped directly to one of the domain ontologies, e.g. databoxes (see below) to the Privacy Preference Ontology (PPO) (Sacco & Breslin, 2012), or personal profiles to the Nepomuk Contact Ontology[9] (NCO). Therefore, a major role of the UI is to abstract the complex ontological knowledge under a simple and intuitive tool that can still enable users to take full advantage of the semantic core. In addition, the di.me UI also allows people to personalize some of the ontologies across the UI. For example, profiles can be customized by allowing users to extend NCO attributes and personalize how they describe themselves as an individual.

Objects for Personal Information Items. Personal information in the di.me domains cover two distinct types of concepts: Concepts that refer to personal information which identifies and describes the user as an individual; and concepts for user item collections, consisting of personal information items (such as documents, images, etc.). In the object model, this distinction is reflected in by the following concepts:

- *Profile card:* As central UI object representing personal identities the term "profile card" together with a card-visualization is used. It contains arbitrary numbers and selections of attributes describing the user, such as names, forms of contact, images, birth date, etc. Attributes are also modeled objects, as the user shall be able to re-use in several profile cards. Multiple profile cards may be used, potentially with different pseudonyms and non-overlapping sets of attributes. In the main navigation, the list of profile cards is labeled with "My Profile".

  Each profile card is either associated with information synchronized from a particular service (e.g. personal profile information from LinkedIn), or can be set up and extended freely by the user.

[9] `http://www.semanticdesktop.org/ontologies/nco/`

- *Databox:* A "databox" refers to a virtual collection of common information items that are typically shared by users. Databoxes can be used to categorize information items, and customized databoxes for individuals and groups can be created.
- *Data:* "Data" refers to a single informational item coming from the user's personal information sphere that can be identified in a file. This includes documents (e.g. PDF, doc, xls) or images.

Test results. In the usability tests, these concepts and terms were presented and discussed with the test participants. While no particular issues arose for 'profile card' and their attributes, the original wording "Information Sphere" for the list of data had to be changed because it was not self-explaining. It was renamed to the term "My Data" which, however, in the second usability study still was confused with the profile cards by some test participants. A suggested optimization was "My documents".

The term "databox" was self-explaining for most test participants. However, based on the information provided in the tests, the uni-directional sharing in databoxes was not clearly evident to half of the participants.

Objects for the Social Sphere. The following pair of UI objects describes the social sphere:

- *Person:* The object "person" refers to a contact of the user. Depending on the use scenario, the personal di.me nodes will retrieve persons from already existing databases, e.g. by synchronizing the user's personal data (e.g. email contacts, or contacts known through a service). The term "People" is additionally used as label in the UI for the main view.
- *Group:* Groups can either refer to a number of known persons manually defined by the user, a group of persons associated to an external service, or a group of persons suggested by the system depending on the current context (e.g. a group of nearby-people). A person can be assigned to many groups. Whether the popular concept "group" is sufficient to structure a user's social sphere appropriately was an important design question. For the di.me userware, the concept should cover as well fine-grained sets of persons, e.g. the attendees of a meeting, groups like "friends", but also high-level concepts like "private lifesphere".

These objects form not only the system's information structure but also organize the access rights management: Databoxes, data, and profile cards can be configured to be shared with groups or individual persons.

Test results. As expected, these terms were self-explaining to the users. When asked how they would organize personal groups, few test participants suggested hierarchical groups for higher level groups like "business". The statement of most participants, however, was that they would use only few groups like "friends", "colleagues" etc. Only in some cases would assign people to multiple groups.

The service-provided groups (shown by a group of Facebook-friends in the test scenarios) were not expected together with the other groups by half of the participants.

These service-provided groups should be clearly identifiable as such and also be accessible via a list of connected services.

3.2 View Structure and Navigation Concept

The view structure has to visualize the UI objects and their relations in a flexible navigation. For that, the design approach of an object-oriented UI structure was adopted: Navigation and views are designed along the data UI objects, the relations amongst them, their attributes etc.

As *navigation concept for main views,* a typical tabbed view was chosen for the web UI (figure 1). The tabs form the main navigation for the main views and are offered at the top of the di.me use interface as a fixed header.

A first main view "People" (selected in figure 1) combines lists for the objects people and groups. Similarly, the data objects and the databoxes are shown in the combined view "My Data". This view type shows the full list of all elements on the top, and the expandable list of categories (groups or databoxes) in the bottom. The layout allows for a drag-and-drop interaction for assigning elements to categories. A second view type shows flat lists of elements, e.g. as thumbnails. This type is used for the main view "My Profile", which displays the different profile cards. Further main views were offered for messaging, system settings, or—only in some of the tested concepts—specific visualizations (a timeline, a configurable dashboard).

One central UI element that is being shown in the main views is the *properties bar* (figure 1, right margin column). It shows attributes of the currently selected object, general information (e.g. file name, file size, file path), but also the di.me access rights, related to the awareness of shared information (e.g. a list of files or databoxes that a person has access to). In addition, the properties bar can contain shortcuts to certain actions related to the object at hand.

For details of an element, different *sub-pages for object details* were developed. In addition to the page-oriented views, *modal dialogues* are used for actions like assigning items (e.g. data) to a category-object (e.g. databox), or sharing information to a person or a group via a particular identity (see section 4.3).

Test results. Within the first versions of the concept, two views for the *data and databoxes* were offered, whereby the view type for 'data" was focused on flexible searching and filtering, and the view "databoxes" on configuration of access rights. This separation, however, was criticized by the majority of the test participants. In fact, many participants spontaneously recognized the similar relation of people to groups and data to databoxes. Therefore they suggested similar layouts for both views.

Some insights were also gathered on the *priority of views*: The test particapants regarded the connection to different external services as one major functionality because this supports an easy overview of contacts and data in different systems and storages. For that, for users experienced with multiple communities the most important view was the list of people and groups, together with a direct visualization who is known via which service.

The *properties bar* (figure 1, right column) as common element for the main views was introduced based on the first usability test. Test participants strongly demanded a stable place for information about which information is accessed by a person, who has access to what personal information item by which service. Again, in particular power users working with multiple systems stressed this advantage. Also the requirement of getting a quick overview on people or groups can be realized if the properties bar shows aggregated information on multiple selected persons, groups, or data items.

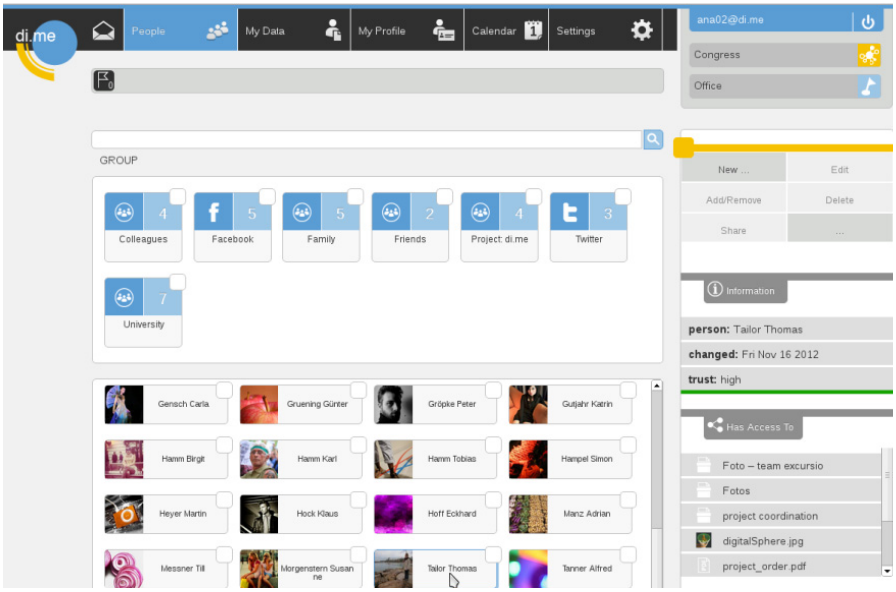
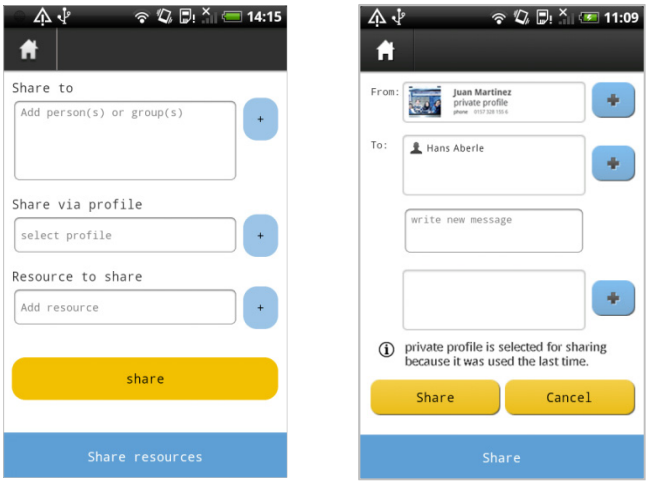


Fig. 1. Web UI: technical prototype with optimizations after the usability tests

3.3 Sharing via Multiple Identities

When shaping the presentation and interaction with identities in the di.me userware, the design decision was taken to offer *only* the profile cards for selecting the partial identity if it is needed for communicating or sharing with others. This implies that the non-global identifier, the service-account, is linked to the profile card: Selecting a profile card implicitly selects the service-account. As a result, the selection of a profile card decides on linkability and unlinkability: All information sent via a profile card is linkable to this identity. Other options were considered, e.g. to bind the service-account to a group. However, this would require the user to select service-accounts for groups, to define groups by considering identities and linkability for them, and to manage sharing rights for profile cards in addition. With the chosen concept, the service-account is not explicitly shown in the UI and the user only has to work with profile cards.

The selection of a profile card is shown in sharing dialogues (see figure 2a for the mobile client version): To share e.g. a photo, the user chooses an action “Share to ...”. In an opened dialogue, the recipients and profile card to share ‘via” can be selected.



(a) *(b)*

Fig. 2. Dialogues for the mobile UI: (a) “share resource” in the tested functional UI, (b) “Share” in a scribble with optimizations from the second usability test

Test results. In the second usability test, this interaction sequence and the underlying considerations were tested and discussed in detail with the participants. The general metaphor to change the identity used for a communication action by selecting the profile card was easily understood and not questioned by the participants. Having tried out the interaction flow, the participants took for granted that shared information would be linkable to the profile and to other information shared via the same profile.

Few participants mentioned that the content of a selected profile card must easily be checked so that the user can be sure what is being shared. In a wireframe done after the test (figure 2b) a preview of the profile card’s information is added for that.

While the role of the service-accounts as non-global identifiers could not be deeply discussed with most participants, the participants assumed that the recipient cannot link this communication to other identities.

A further aspect was partly difficult to understand: The majority of participants suspected that a message was sent to Facebook if a profile card was selected that is synchronized with a Facebook-Profile. In the interview, the participants explained that the wording “via” (see figure 2, a) suggested this. Even if such a functionality is planned, the wording should point more to the fact that a sender-identity is selected. So, this could be changed to a label “From” as used for email-accounts (scribble in figure 2 b). As soon the function of pushing content via third services will be in place, this may additionally be indicated in the UI.

4 Summary

As an important result of the user-centered design, the information and UI structure was developed which represents the system domains of personal information, identities, and social sphere. After the user testing, the key UI concepts and view structure can be considered as stable.

However, further investigations addressing the interplay of the information structure and access rights are necessary: Specific cases like adding a person to a group will potentially change access rights (e.g. if the group has access rights the person did not own so far). For such cases, the awareness of the user must be ensured which requires additional information in the UI. An approach here is developed by the management of contact's trust in the di.me network (Heupel et al., in press).

The results of the user studies point into the direction that the profile cards connected to service-accounts as communication identifiers worked well with respect to the understanding of unlinkable identities. However, some implications of this design approach were not deeply investigated so far. E.g., if information is shared to a group, different profile cards may be known to its members. If a single card is chosen for sharing, it would be newly disclosed to a part of the group members. In this case, a *linkability warning* could be presented. A further requirement not investigated in the studies is the case of *intended linkability:* The user may want to inform a recipient that two profile cards are coming from the same sender. While users always may inform the recipient 'manually', the system should also support linking two profile cards.

Acknowledgments. This work is funded by the European Commission under the Seventh Framework Program FP7/2007-2013 (digital.me – ICT-257787).

References

1. Al-Sinani, H.S., Alrodhan, W.A., Mitchell, C.J.: CardSpace-liberty integration for CardSpace users. In: IDTRUST 2010, Proceedings of the 9th Symposium on Identity and Trust on the Internet, pp. 12–25. ACM, New York (2010)
2. Crumlish, C., Malone, E.: Designing Social Interfaces: Principles, Patterns, and Practices Improving the User Experience. O'Reilly, Sebastopol (2009)
3. Dumas, J.: User-based evaluations. In: Jacko, J.A., Sears, A. (eds.) The Human-Computer Interaction Handbook, pp. 1093–1117. Lawrence Erlbaum Associates, Mahwah (2003)
4. Heupel, M., Bourimi, M., Kesdoğan, D.: Trust and Privacy in the di.me Userware. In: Kurosu, M. (ed.) Human-Computer Interaction, Part III, HCII 2013. LNCS, vol. 8006, pp. 39–48. Springer, Heidelberg (2013)
5. Heupel, M., Fischer, L., Kesdogan, D., Bourimi, M., Scerri, S., Hermann, F., Giménez, R.: Towards robust multidimensional trust for the digital.me userware by considering situational context. In: 2012 5th New Technologies, Mobility and Security (NTMS), pp. 1–6 (2012)
6. Lewis, J.R.: Usability Testing. In: Salvendy, G. (ed.) Handbook of Human Factors and Ergonomics, 4th edn. Wiley (2012)

7. Pfitzmann, A., Hansen, M.: A terminology for talking about privacy by data minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management (Version v0.34) (2010), http://dud.inf.tu-dresden.de/Anon_Terminology.shtml (retrieved)
8. Sacco, O., Breslin, J.G.: PPO & PPM2.0: Extending the Privacy Preference Framework to provide finer-grained access control for the Web of Data. In: I-Semantics 2012 - 8th International Conference on Semantic Systems (2012)
9. Scerri, S., Schuller, A., Rivera, I., Attard, J., Debattista, J., Valla, M., Hermann, F., Handschuh, S.: Interacting with a context-aware personal information sharing system. In: Kurosu, M. (ed.) Human-Computer Interaction, Part V, HCII 2013. LNCS, vol. 8008, pp. 122–131. Springer, Heidelberg (2013)
10. Schwarte, P., Bourimi, M., Heupel, M., Kesdogan, D., Giménez, R., Wrobel, S., Thiel, S.: Multilaterally secure communication anonymity in decentralized social networking. In: Proceeding of 10th International Conference on Information Technology: New Generations, ITNG (in press, 2013)
11. Scerri, S., Gimenez, R., Hermann, F., Bourimi, M., Thiel, S.: digital.me – Towards an integrated Personal Information Sphere. In: Proceedings on the Federated Social Web Summit 2011 (2011)
12. Scerri, S., Cortis, K., Rivera, I., Hermann, F., Bourimi, M.: di.me: Context-Aware, Privacy-Sensitive Management of the Integrated Personal Information Sphere. In: EU Project Networking Track, 9th Extended Semantic Web Conference, ESWC 2012 (2012), http://data.semanticweb.org/conference/eswc/2012/paper/project-networking/372 (retrieved)
13. Schümmer, Lukosch: Patterns for Computer-Mediated Interaction. John Wiley & Sons, Chichester (2007)
14. Simonite, T.: Google Invites Everyone to Be Friends on Its Social Network (2011), http://www.technologyreview.com/view/506796/algorithm-predicts-circles-of-friends-using-contacts-data/ (retrieved)
15. Steuer, J.K., Fernando, R., Bertino, E.: Privacy preserving identity attribute verification in windows cardspace. In: DIM 2010, Proceedings of the 6th ACM Workshop on Digital identity Management, pp. 13–16. ACM, New York (2010)
16. Sun, S.-T., Pospisil, E., Muslukhov, I., Dindar, N., Hawkey, K., Beznosov, K.: OpenID-enabled browser: towards usable and secure web single sign-on. In: CHI EA 2011, CHI 2011 Extended Abstracts on Human Factors in Computing Systems, pp. 1291–1296. ACM, New York (2011)
17. Thiel, S., Bourimi, M., Giménez, R., Scerri, S., Schuller, A., Valla, M., Fra, C., Wrobel, S., Hermann, F.: A Requirements-Driven Approach Towards Decentralized Social Networks. In: Park, J.J., Leung, V.C., Wang, C.-L., Shon, T. (eds.) Future Information Technology, Application, and Service. LNEE, vol. 164, pp. 709–718. Springer, Heidelberg (2012)

Trust and Privacy in the di.me Userware

Marcel Heupel, Mohamed Bourimi, and Doğan Kesdoğan

Institute for Information Systems, University of Siegen, Germany
heupel@wiwi.uni-siegen.de

Abstract. People in online social networks are constantly sharing information with different audiences and it becomes more and more difficult for them to keep track of distributed information. Further, due to the complexity of the digital landscape, it is a constant risk to unintentionally share information to the wrong people, possibly leading to a loss of reputation. The European research project di.me is concerned with the development of a userware, empowering end-users to keep track of their digital footprints, with an intelligent user interfaces (UI) and smart advisory. In this paper we present how we calculate persons trust and inform the privacy of resources shared among persons. We show the concepts for trust and privacy advisory in the di.me userware and address problems, we were confronted within the design and evaluation process and how we tackled them. In this respect we specifically address change requirements (i.e. trust model and UI improvements) we conducted after an evaluation and user trials with a first prototype.

Keywords: Trust, privacy, user interface, privacy advisory, online social networks, di.me.

1 Introduction

In todays online social life people are constantly sharing information. It is not easy for end-users to keep track of all their information, distributed over different online social networks (OSNs). When in the beginning of OSNs some years ago most information posted was public, the privacy awareness of people raised over the last years. Nowadays, for most applications it is common practice to configure the visibility of posted information with the help of security settings or privacy preferences in the user interface (UI).

The strong interplay of security and usability, more precisely, the fact that usability is an important prerequisite for secure systems, is getting more and more attention recently (see e.g., [1]) and it was already mentioned by Kerckhoff in 1883 [2] in his six design principles for military ciphers. This is even more crucial as user experience (UX) and usability have also consequences for privacy and trust, as it is well known that usability and UX are important factors for trust [3]. A recent study of Madejski et al. [4] reported, almost all people make errors when setting their privacy preferences in OSNs. This leads in the end to information being shared with wrong audiences without intention and can cause serious damage to the personal reputation in extreme cases. The European

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 39–48, 2013.

research project di.me is developing a tool, integrating all information from connected OSNs in a single-user controlled information sphere. Key functionalities in this respect are the provision of an privacy-preserving intelligent UI and the provision of smart advisory.

In this paper we will present our approach providing smart trust and privacy advisory when sharing information in OSNs. The approach consists of a trust metric, bringing together the privacy of information and trust of persons, which was already presented in [5] and several UI concepts in order to adequately present the advisory (based on that metric) to the user in an intuitive and non-intrusive way. The first version of the metric, as well as corresponding UI concepts, has been evaluated in first prototypes of the di.me userware. Main focus of this paper is the presentation of the current status of work as well as addressed change requirements (CRs) we identified in the evaluation.

The paper is structured as follows: In Section 2 we elaborate key concepts for trust and privacy in di.me and put them in relation to related work. In Section 3 we provide some essential background information about the di.me userware. In Section 4 we present our identified CRs, which we address in our approach, discussed in Section 5. Finally we conclude the paper in Section 6.

2 Definition of Trust in di.me and Related Work

Trust and privacy advisory is one of the key functionalities of the di.me userware, so it was of major importance to find a common understanding for those terms, and to develop adequate design concepts for the UI to prevent misunderstandings on the side of the end-users. Since it is not easy to give a short definition of trust, it is even harder to present this in an intuitive way to the user. The concept of trust is difficult to perceive, and the vast amount of definitions make it even harder. Therefore, we will expound here the concept of trust as we use it in di.me. In the past, there has been a lot of research in the field of trust, and there are also a lot of different definitions. When talking about trust, many people think about trust establishment with unknown parties (e.g., [6] and many more). This is strongly related to reputation, since to the decision to trust an entity or not is based on the public opinion. There are a lot of papers concerned with trust in social networks. Usually they talk about trust in the OSN provider itself or about trust in unknown entities (e.g., when adding new contacts etc.), which is also more going in the direction of reputation (*to trust someome because of good reputation*, see also [7]).

The di.me userware deals with user-centric information, contact management, and it supports the user by interacting with already existing contacts in different OSNs. Since the establishment of new connections is of minor relevance in the context of online social networks[1]. Therefore, there is no concept of reputation in di.me, not even a concept for friends-of-friends [8]. Consequently the concept of trust in di.me differs from the previously mentioned definitions. In general the

[1] We assume most people do not consider reputation to decide if they connect to a person in a social network or not.

di.me concept of trust can be defined like the following, which is very similar to Josangs [7] definition of *Reliability trust*. To trust someone means that the other party behaves as expected with a specific probability. This is of course always connected to a certain context. In di.me we are more or less only moving in the context of information disclosure, which means basically that trusted persons will give private information with a very low probability to (untrusted) third parties. Trust in di.me tries to measure the personal direct trust of the user to each of his contacts. It is formed by the interaction between individuals (e.g., communication, sharing) and can consequently also be seen as measure for the strength of relationship. Sharing private information to a person is an expression of trust. As described in [5], we compute the trust based on previous interactions, especially the disclosure of information. The more private the disclosed information is, and the smaller the audience, the higher is the resulting trust value. Related to privacy in OSNs many works deal with supporting the user in privacy preferences (see e.g., [4,9,10]), which are rather static. Our approach, of giving dynamic trust based privacy advisory on runtime, is rather novel and, to our best knowledge, not covered by other works.

3 Background Information

3.1 The di.me Project

The di.me project targets the integration of personal information in a personal information sphere, in order to give users full control over their data and allow intelligent support and advisory [11]. The di.me userware operates on user-controlled servers, the so called *di.me Personal Servers* (PSs) and can be accessed either by a Web-based UI or an Android application. It aggregates information, gathered from connected services, like social networks or special services for e.g, processing location information. The PSs' form a decentralized network and also implement State-Of-The-Art network security and anonymity mechanisms, like e.g., Tor, StartTLS or OAuth in order to ensure maximum protection of users' privacy and their data.

Collected information is stored in a semantic repository enabling intelligent advisory and advanced techniques like e.g., semantic matching of profiles [12].

3.2 The di.me TrustEngine

The component in the di.me userware, processing trust and privacy in order to calculate respective advisory, is called *TrustEngine*. It is located on the PS and accesses information stored in the semantic core as well as in live context data in order to trigger warnings and advisory, shown in the UI. The TrustEngine is integrated in all interaction flows concerned with the manipulation and especially the disclosure of information (e.g., share a profile to a group, send a file to a contact, etc.). Figure 1 depicts a simplified architecture of the di.me userware in order to illustrate the integration of the TrustEngine into the core of the

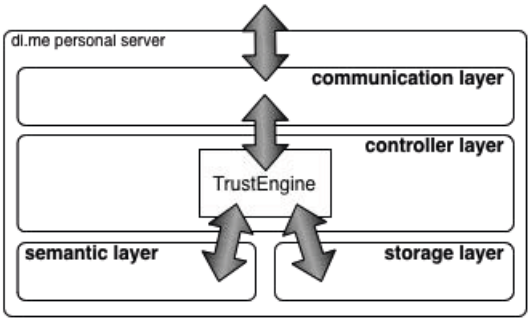


Fig. 1. The TrustEngine in the architecture of the di.me userware

system. The core concept to calculate trust and trigger privacy warnings and advisory was already presented in Heupel et al. 2012 [5]. The general idea is that all information stored within the PS should be classified regarding the privacy of the content. To classify information we us a scale from 0 to 1, where 0 is considered public information and 1 secret[2] The trust value for contacts can then be calculated based on the information shared with them. The disclosure of very private information to a person is an implicit expression of trust. This approach would of course require a short learning phase before it has enough confidence to detect privacy flaws. Once it is calibrated, it is easy to detect if private information is shared to untrusted persons. Further it can detect a possibly wrong classification of information when information marked as private, is shared to a lot of people, it is very likely that the classification is wrong. Besides the automatic calculation of trust values, it is also possible to set it manually in the UI, overriding the calculated value.

3.3 First Prototype and Evaluation

The first prototype was presented to the visitors of a summer school event in Segovia in July 2012. The visitors had the possibility to test the application on a booth or even download it to their own Android phones and were asked to fill out a questionnaire. Additionally a focus group was selected, which got a special presentation and tried the di.me userware on several provided devices. Figure 4 shows screenshots of the first prototype.Most interesting in the evaluation of the first prototype, concerning trust and privacy advisory, was to get an impression if the test users understood the general concept, if they thought it was something useful and if the concepts are presented in an intuitive and non-intrusive way in the UI[3]. Especially the members of the focus group (12 students with

[2] A value of 1 should here really be understood as top secret, e.g., like a private encryption key, which is usually never communicated to anyone.

[3] The interested reader can find more information about the overall UI concepts in [13].

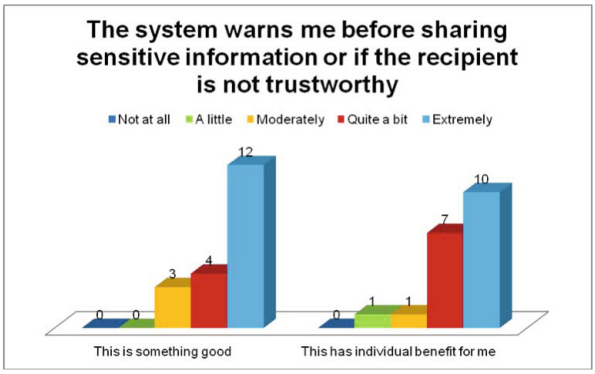


Fig. 2. Answers to one of the questions about privacy advisory in the questionnaire

technological majors) were really interested in the feature of privacy advisory and considered it an extremely useful feature (see also Fig. 2 for some of the evaluation results).

4 Requirements Gathering

Besides the general usefulness of the features, we asked also more concrete questions about selected features in the UI, if they are usable and understandable. By doing this we identified several issues in our concepts[4], from which we address the main critical ones in the context of this paper. In the following we will shortly discuss the change requirements (CR1-CR3) we deducted from those issues.

A major critical point of the first prototype was concerned with the general trust model. There is only one trust value for each person and only only privacy value for each sharable information item. This model turned out to be not covering all different use-cases. The di.me userware is intended to cover all different lifespheres of the user, this means business as well as private. We identified that trust and privacy can be interpreted slightly different, depending on the lifesphere[5]. Especially when dealing with intersection of different lifespheres (like e.g., business and private) this can lead to situation where an unintended information disclosure is not detected. An example for such a situation could be: The user shares accidentally a highly confidential document to his grandmother. Obviously this case will not be detected. Therefore we need to extend the model to detect information being shared to a context, it does not belong to (CR1).

The second important issue we identified, was concerned with the presentation of the trust and privacy values in the UI and how they can be manipulated. The presentation of trust and privacy values in the UI is no trivial task, since it can be easily misinterpreted. As shown in Fig. 4, the privacy value of files, as well

[4] Developed internally by involving selected di.me consortium partners.

[5] The privacy value of confidential business documents can mean something else than for private photos.

(a) Privacy slider (b) Trust slider

Fig. 3. Screenshots of trust and privacy slider in the first prototype

as the trust value for persons are represented as a continuous slider. The label next to the slider was showing the value as a percentage value, depending on the position. This turned out not to be very intuitive. The users were not sure what exactly the setting to a high or low trust means, respectively for the privacy value. It turned out to be easy to misunderstand, if a low trust in someone would mean just *no trust* (e.g., due to a lack of information) or *distrust* (meaning to expect someone to behave to the users' disadvantage). Therefore we needed to redesign the UI elements representing trust and privacy values, in order to make the concept really clear to the user (CR2).

The calculation of trust advisory takes place on the server and is triggered by all calls involving manipulation of data and access rights. Therefore, the TrustEngine is involved if information is shared to a person or group, but also when adding persons to a group or documents to a databox. For all those actions a HTTP request to the PS is necessary, which can produce a lot of communication overhead and lead to a bad UX in the end. Therefore the third identified requirement is to optimize the communication flow for trust advisory (CR3).

5 Approach

5.1 Extending the TrustEngine (CR1)

One of the main critical points that has been identified in the evaluation of the first prototype was, that the trust metric we used might not be applicable in all use-cases, especially when dealing with intersection of different lifespheres (e.g.,

business and private). To overcome this problem, we analyzed the problem and extended our approach from [5] respectively. Our solution envisages to use the groups[6] in the systems to identify lifesphere borders[7].

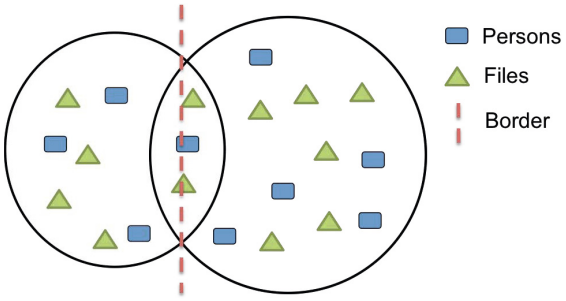


Fig. 4. Illustration of two overlapping groups with common files and persons

We assume that users usually tend to group the contacts and thereby automatically align them to their lifespheres. To foster this behavior we decided to have 3 default groups in the system, for family, friends and business contacts. On top of this, there are automatic created groups for contacts, imported from existing social networks (e.g., all LinkedIn contacts) and automatic ad-hoc groups that are created based on location context information, if contacts are nearby for a certain time (e.g., sitting in the same room during a meeting). In order to identify information passing over lifesphere borders without the users intention, the first step is the identification of such borders. Naturally not every group represents a separate lifesphere (Fig. 4 illustrates two overlapping groups with some common files and contacts). In oder to identify a possible threat, we compute the pairwise distance of two groups, the one the information is already related to[8] and the recipient group. Information becomes "attached" to a group by being shared to members of this group.

In order to calculate the distance between two groups, we consider the two most relevant factors: common contacts and common files. Since we are interested in the difference, and not what both groups have in common, we take the symmetric difference ($\triangle$), of both sets (defined for sets Y, Z in (1)). In a coordinate system both values are orthogonal to each other, so the distance between two groups (= distance to the origin) can be calculated like shown in (2).

$$Y \triangle Z = (Y \setminus Z) \cup (Y \setminus Z) \tag{1}$$

[6] A group in di.me is only for ordering contacts on user side, and should not be confused with "discussion groups". They can be compared to the "circles" of Google+ or the user defined friend lists of Facebook.

[7] A border in this context e.g., between business and private lifesphere is weak in most cases, since there can be colleagues who are also friends or family.

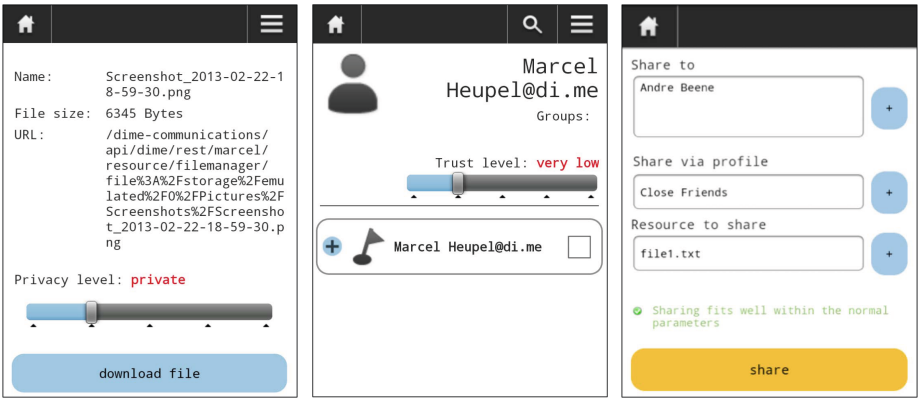
[8] Information becomes related to a group, by being shared to members of those groups before.

$$GD = \sqrt{\left(\frac{|A \triangle B|}{|A \cup B|}\right)^2 + \left(\frac{|C \triangle D|}{|C \cup D|}\right)^2} \tag{2}$$

With this mechanism we can calculate a distance between two groups, where $\sqrt{2}$ is resembling the maximum distance (both groups are completely disjunct). Combined with the privacy value, the new distance can now be used to trigger additional warnings, not covered by the previous approach.

5.2 Reworking UI Concepts (CR2)

As a representation of the trust value in the first prototype, we used a continuous slider and labels showing a percentage value depending on the position. This turned out not to be very intuitive. The test users were not sure what exactly the setting of a trust of e.g., 50% means and the same for privacy. Therefore we consolidated the user feedback and reworked our UI concepts together with UI design experts. Figure 5 (a, b) shows some screenshots of the improved UI for the Android prototype. We left the continuous slider, but the label is no longer showing a percentage value. Instead we use simple text labels like "private" or "public" for the privacy of files and "trusted", "untrusted" for trust. Further, the labels are changing their color depending on the restrictiveness (red→private/untrusted, yellow→medium trust/privacy, green→public/trusted). To reduce the arbitrariness of setted values (e.g., how should someone decide if a file is 64% or 67% private?), we introduced 3 fixed steps on each slider (left, middle, right) and thereby drastically reduced complexity and uncertainty of the user setting the values.



(a) Privacy slider (b) Trust slider (c) Trust awareness

Fig. 5. New trust and privacy concepts in the Android UI

5.3 Optimizing the Interaction Flow (CR3)

In the first version of the clients the TrustEngine was completely integrated in the server, checking all manipulations of data for possible trust and privacy issues. If such an issue was detected, a warning will be sent to the UI instead of an acknowledgment of a successful performed manipulation. The user had then different options to dissolve the issue (e.g., by removing untrusted persons from the recipients) or to ignore the warning and share anyway, which would lead to an adaption of trust values of the recipients (like described before). This approach did work well, but in practical tests we discovered some potential to be optimized as we identified in section 4. Therefore we changed the API a bit, and moved the logic of advisory calculation for file sharing to the client. By doing this we were able to improve the user experience due to a recognizable reduced communication overhead. We were also able to provide additional awareness features in the UI, like e.g., a realtime indicator showing if there is a possible trust issue when adding people to a group of recipients. Figure 5(c) shows two of the new colored textfields, indicating the trust status.

6 Conclusions

In this paper we presented how trust and privacy are addressed in di.me and discussed selected issues identified in first evaluations and user trials of the di.me userware related to them. We also presented improvements to the general model, the UI and the interaction design in order to solve those issues. Thereby we were able to improve the warning mechanism to identify information, shared with the wrong audience, made the UI more intuitive and increased the response time for showing privacy advisory in the client UI.

Since the di.me project follows an agile approach merging best-practices from the security, usability and HCI community [14], we will continuously conduct further evaluations and improve the system accordingly. Especially the new concepts introduced in this paper will be analyzed and evaluated again, with end-users and experts. Besides this, we will further extend the privacy advisory by including analysis of microposts with NLP and including live context information going a step further as we already proposed in [15]. Another target for future work is the improvement of initialization of trust and privacy values in order to reduce the learning phase as well as the need for manual settings to a minimum.

Acknowledgements. The work carried out in order to write this paper was supported by the Seventh Framework Program of the European Union, (FP7/ 2007- 2013), in the digital.me project under grant agreement no. 257787. Special thanks goes to our consortium partners from Fraunhofer IAO for their valuable contributions.

References

1. Cranor, L.F., Garfunkel, S.: Security and Usability: Designing Secure Systems That People Can Use. O'Reiley (2005)
2. Kerckhoffs, A.: La cryptographie militaire. Journal des Sciences Militaires IX, 5–38 (1883)
3. Corritore, C.L., Kracher, B., Wiedenbeck, S.: On-line trust: concepts, evolving themes, a model. International Journal of Human-Computer Studies 58(6), 737–758 (2003)
4. Madejski, M., Johnson, M., Bellovin, S.M.: A study of privacy settings errors in an online social network. In: International Conference on Pervasive Computing and Communications Workshops (PERCOM), pp. 340–345. IEEE (2012)
5. Heupel, M., Fischer, L., Kesdogan, D., Bourimi, M., Scerri, S., Hermann, F., Gimenez, R.: Context-Aware, Trust-Based Access Control for the di.me Userware. In: 5th International Conference on New Technologies, Mobility and Security (NTMS 2012), pp. 1–6 (2012)
6. Aberer, K., Despotovic, Z.: Managing trust in a peer-2-peer information system. In: Proceedings of the Tenth International Conference on Information and Knowledge Management, pp. 310–317. ACM (2001)
7. Jøsang, A., Ismail, R., Boyd, C.: A survey of trust and reputation systems for online service provision. Decision Support Systems 43(2), 618–644 (2007)
8. Heupel, M., Bourimi, M., Scerri, S., Kesdogan, D.: Privacy-preserving concepts for supporting recommendations in decentralized OSNs. In: Proceedings of the 4th International Workshop on Modeling Social Media, MSM 2013. ACM, New York (to appear, 2013)
9. Fang, L., Kim, H., LeFevre, K., Tami, A.: A privacy recommendation wizard for users of social networking sites. In: Proceedings of the 17th ACM Conference on Computer and Communications Security, pp. 630–632. ACM (2010)
10. Chaytor, R., Brown, E., Wareham, T.: Privacy advisors for personal information management. In: SIGIR Workshop on Personal Information Management, Seattle, Washington, pp. 28–31 (2006)
11. Thiel, S., Bourimi, M., Giménez, R., Scerri, S., Schuller, A., Valla, M., Wrobel, S., Frà, C., Herman, F.: A requirements-driven approach towards decentralized social networks. In: Proceedings of the International Workshop on Social Computing, Network, and Services (2012)
12. Cortis, K., Scerri, S., Rivera, I., Handschuh, S.: Discovering semantic equivalence of people behind online profiles. In: Proceedings of the 5th International Workshop on Resource Discovery, RED 2012 (2012)
13. Hermann, F., Schuller, A., Thiel, S., Knecht, C., Scerri, S.: The di.me user interface: Concepts for sharing personal information via multiple identities in a decentralized social network. In: Kurosu, M. (ed.) Human-Computer Interaction, Part I, HCII 2013. LNCS, vol. 8006, pp. 29–38. Springer, Heidelberg (2013)
14. Bourimi, M., Barth, T., Haake, J.M., Ueberschär, B., Kesdogan, D.: AFFINE for enforcing earlier consideration of nFRs and human factors when building socio-technical systems following agile methodologies. In: Forbrig, P. (ed.) HCSE 2010. LNCS, vol. 6409, pp. 182–189. Springer, Heidelberg (2010)
15. Bourimi, M., Rivera, I., Scerri, S., Heupel, M., Cortis, K., Thiel, S.: Integrating multi-source user data to enhance privacy in social interaction. In: Proceedings of the 13th International Conference on Interaccion Persona-Ordenador, INTERACCION 2012, pp. 51:1–51:7. ACM, New York (2012)

From Living Space to Urban Quarter: Acceptance of ICT Monitoring Solutions in an Ageing Society

Simon Himmel, Martina Ziefle, and Katrin Arning

Human-Computer-Interaction Center, RWTH Aachen University, Germany
himmel@comm.rwth-aachen.de

Abstract. The question of designing safe, secure, and human-centered urban living environments is complex, as different and controversial needs from different sources have to be harmonized. The ongoing demographic change with more and more older and frail persons living in urban environments raise the question how modern technologies in the information and communication sector can be seamlessly integrated in specific urban spaces. Not only (medical) safety and data security issues need to be considered but also technology acceptance and the fragile trade-off between the ubiquity of technologies on the one and the perceived trust in and reliability of technologies on the other hand are of pivotal importance. To understand users' willingness to accept monitoring technologies in different locations (ranging from home to public spaces) we pursued an exploratory approach. 127 participants in a wide age range volunteered to take part in the questionnaire study. Results show that individuals independently of age and gender are quite reluctant to accept such technologies in general, even for medical reasons. It was ascertained that acceptance was impacted by the type of technology and the space (domestic vs. public) and the health status of the people.

Keywords: Technology Acceptance, ICT Health Monitoring, Living Spaces, Urban Quarter, User Diversity.

1 Introduction

As a matter of fact the demographic change in most (western) societies raises considerable challenges for urban environments with respect to providing humane and human-centered living conditions. Characteristically, the requirements for designing safe, secure, and human-centered urban living environments are highly complex. Different and sometimes controversial needs have to be taken into account. These requirements include the harmonization of mobility aspects, the integration of safety and security of (medical) technologies in the context of treatment and care up giving sensitive consideration to the perceived intimacy, privacy and data security issues as well as individuals' fears of losing control. In addition, as communities might differ regarding their economic wealth, also cost burden need to be respected. Politics and communities take up the challenge and put effort into expanding conventional and well-established solutions by investing in retirement homes and the education of

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 49–58, 2013.

elderly care nurses [1]. For the next half of the decade all efforts cannot provide suitable solutions to the challenge of a numeral increase in older and care-needing persons in contrast to the continuous decrease in younger and potential caring people [2]. While some age-induced issues like dementia or wound treatment require personal assistance, a great number of older adults endangered by sudden breakdowns [3] are at risk because they live alone and might not be found in case of emergency [4]. In these cases different surveillance technologies could provide possible solutions, but acceptance of these technologies, if researched at all, is desperately low.

While there is only restricted knowledge about the acceptance of integrated information and communication technology (ICT) in domestic spaces [5],[6], it is well known that public surveillance (CCTV) is controversially discussed especially because of peoples' fear of data privacy violation and the loss of control over the protection of intimacy [7],[8]. Also, it was found that acceptance for technology differs depending on nationality [9] and culture [10].

So far, research does not sufficiently address the role of context factors that might have a considerable impact on the extent to which persons would be willing to use surveillance technologies in urban and private spaces.

2 Research Aims

The willingness of citizens to accept technologies in domestic and urban environments is a highly complex and also fragile phenomenon, which is impacted by a multitude of environmental, spatial, and individual factors, driven by personal needs and using motives. In order to gain a first insight into this complex interaction system, we pursued an exploratory approach in which we selected the monitoring technology type (camera, microphone and positioning system) within the context of medical monitoring. As technology acceptance for or against the medical monitoring is considerably influenced by the specifics of the location or space in which such technologies would be installed, we also varied different spaces, ranging from home (sleeping and living room, garden) to public spaces (park, station, shopping mall). Understanding the trade-off between the perceived usefulness of technologies in the different locations/spaces and the extent of wished control [11] and need of privacy [12], participants had to evaluate the extent of willingness to use the technical systems in the different locations.

3 Methodology

In this section we describe the procedure of the questionnaire study, the variables and the sample.

3.1 Questionnaire Study

To reach a broad user group, we applied the questionnaire technique, addressing users of different health states, ages and gender. The survey was broadcasted online via the

authors' private networks. The questionnaire was designed to obtain information about (1) demographic data (age, gender, health status) and (2) individual living conditions (rural vs. urban districts, alone vs. family, distance to medical care).

The health condition of the participants was determined by four questions regarding chronic diseases, regularly check-ups, dependency on medical technologies and health care. People negating all questions were defined as healthy, whereas people agreeing to one or more of these questions were defined as ill. Then, participants had to evaluate the extent of acceptance. (3) Different integrated surveillance technologies (microphone, camera, positioning system) were to be evaluated in (4) various domestic private spaces (living room, bedroom, garden) as well as in public, urban spaces (local bar, city council, park, train station, museum, church, bus stop). In order to classify the different locations, living spaces were split in three dimensions: private vs. public, intimate vs. distant, and indoor vs. outdoor (see fig.1).

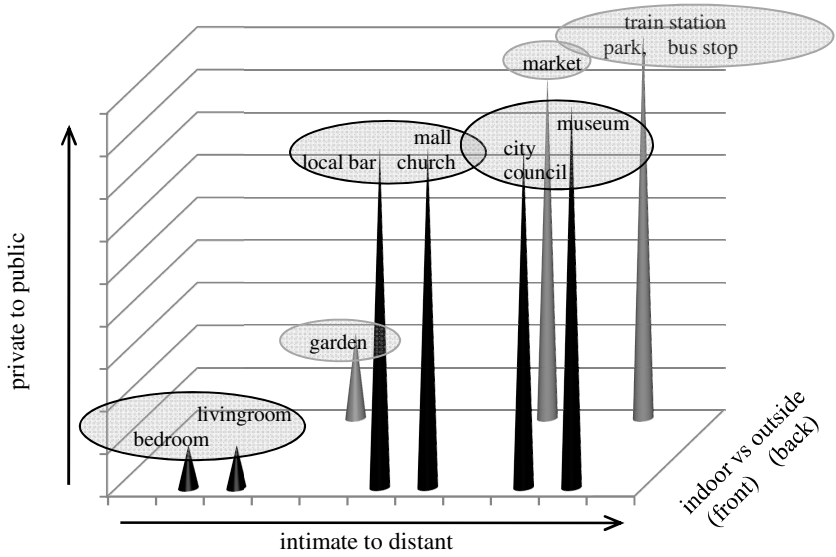


Fig. 1. Classifying regarded spaces in three dimensions

The participants had to judge their willingness to integrate each technology (three) in each space (twelve) on a six point Likert scale (never, no, rather no, rather yes, yes, always).

3.2 Participants

127 participants, aged 19 to 74 years, took part in this study (M=37.8, SD=17.9). 41% of the sample was female, 59% male. To investigate if technology acceptance differs depending on age and technology generation, the sample was split in age groups referring to three different technology generations: The young group was aged between 19 and 28 years, mainly university students (40%, M=22.8 SD=2.6, 49% female/ 51%

male), the middle-aged working generation between 29 and 59 years (30%, M=38.2, SD=9.4, 26% female/ 74% male), and the older adults between 60 and 74 years (30%, M=65.9, SD=4,3, 44% female/ 56% male). The overall educational level (40% university entrance diploma, 47% university degree) lay high above average. 70% of the participants were classified as healthy (see 3.1).

4 Results

In a first exploratory research step we correlated the demographic variables with the acceptance items in order to identify the influence of the regarded user demographics on the acceptance. To get a deeper insight in the acceptance of monitoring for each living space and each technology we followed a two-way deductive approach. We first analyzed the general acceptance of our three chosen surveillance technologies (microphone, camera, positioning system) by summing up the acceptance ratings of this technology in all spaces and then checked for significant effects of user diversity (age, gender, health condition) (see 3.1). Then we analyzed the general acceptance of surveillance in all different living and urban spaces (see 3.2). Due to space limitations only the general acceptance of each technology in each space will be dealt with here. Finally, we focused on the most significant correlations by analyzing how the acceptance of camera surveillance in all spaces is affected by the health condition of the user (see 3.3). Q-Q plots and boxplots indicated that all our scales and items were normally distributed. Data were statistically analyzed by MANOVA procedures and ANOVAs. Significance level was set at 5%.

4.1 Acceptance of Surveillance Technologies: Overall Technology Acceptance

If summarizing the different locations and spaces under study, we find a low acceptance of all regarded technologies (camera, microphone, positioning) with no significant differences between them (see fig. 2).

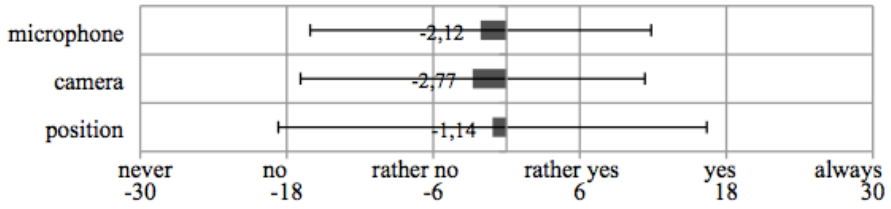


Fig. 2. Mean acceptance ratings and SD for different technology types

Obviously, participants were quite negative about the fact of being monitored at all and no technology was generally accepted, with the lowest levels of acceptance for the camera, followed by the microphone and the positioning system.

Age effects on Technology acceptance. When bringing the age groups into focus, one could see the tendencies of age influencing acceptance differing for each type of technology (see fig. 3). The acceptance of integrated microphones for health reasons marginally increased with age whereas cameras are least accepted by old people, also compared to all other technologies. Camera and positioning system are both most accepted by the middle-aged group. The smallest differences in acceptance were seen with the positioning system showing immunity against the influence of user demographics, which also was a key result in previous research [6]. Though, it should be noted that differences did not reach statistical significance corroborating that age is not driving acceptance: Rather the reluctance of people towards accepting monitoring systems in the medical context is high in all age groups.

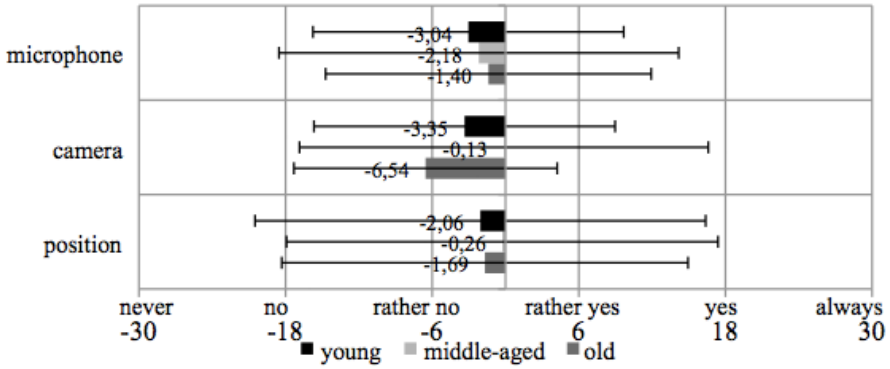


Fig. 3. Mean acceptance with SD by technologies, split by age

Gender effects on Technology acceptance. Different from other studies in which gender did impact the technical self-confidence [13] or interest in specific technologies [14], no significant gender effect could be detected for the acceptance of different technologies for health surveillance purposes (fig. 4).

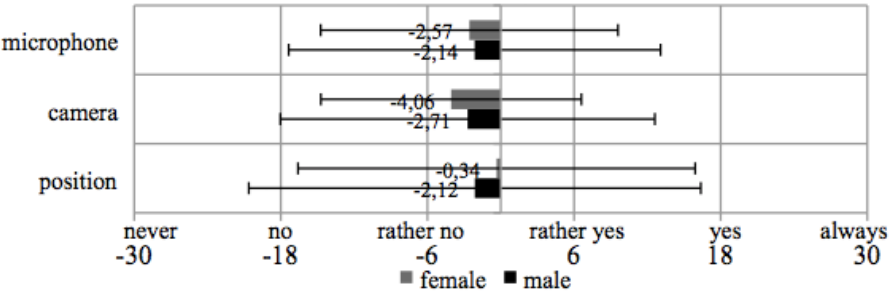


Fig. 4. Mean acceptance with SD by technologies, split by gender

Health condition effects on Technology acceptance. Significant health-effects were revealed for microphone ($F(1,122)=4.24$; $p<0.05$) and camera ($F(1,122)=4.76$; $p<0.05$) surveillance acceptance. Also, the positioning system showed similar but not significant tendencies. Ill people conveyed lower acceptance for health surveillance technologies (see fig. 5).

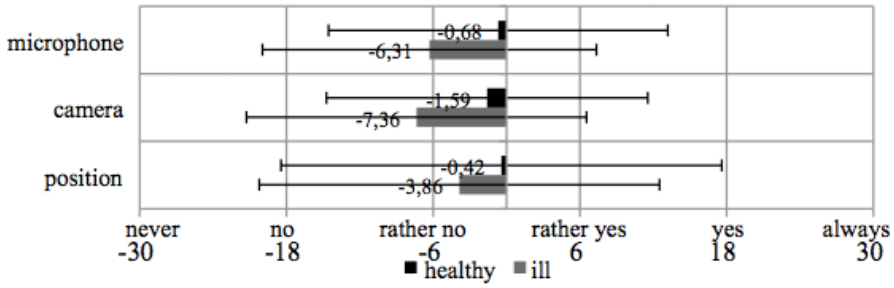


Fig. 5. Mean acceptance with SD by technologies, health-effect

4.2 Acceptance of All Surveillance Technologies in Different Spaces

Breaking down the summarized (general) acceptance of the three technologies into all twelve researched spaces indicated their reciprocal interaction. The consideration of different dimensions for the regarded spaces (see fig.1), seems to depend on the type of technology and needs further research. All data for the whole sample, not considering user diversity, did not show statistical significant effects. The following results are visualized (see fig.6) and discussed descriptively.

For *camera* surveillance the lowest acceptance was recognized when entering private space (living-, bedroom, garden), decreasing from intimate to more distant spaces. This decrease was not transferable to public spaces, where there were no significant differences for private and distant spaces. However, the descriptive data shows positive tendencies for accepting camera surveillance at the train station and bus stop. Regarding all public spaces, a decreasing acceptance of cameras could actually be recognized from indoor to outdoor spaces. The museum as the most accepted place for surveillance plays a particular role – a possible effect might be the acceptance of anti-theft devices. The acceptance of *auditory* surveillance increased from indoor to outdoor spaces, not differing between private vs. public using contexts nor showing different evaluations regarding intimate vs. distant spaces. Microphones' usefulness for surveillance in outdoor spaces could have been regarded as poor for health reasons.

The *positioning system* seems to be independently accepted for all spaces – with exception of the museum, which breaks ranks shouting for further research. It is maybe affected by acceptance of anti-theft motion detecting, whose functionality basically is one kind of positioning system. This result also verifies former research.

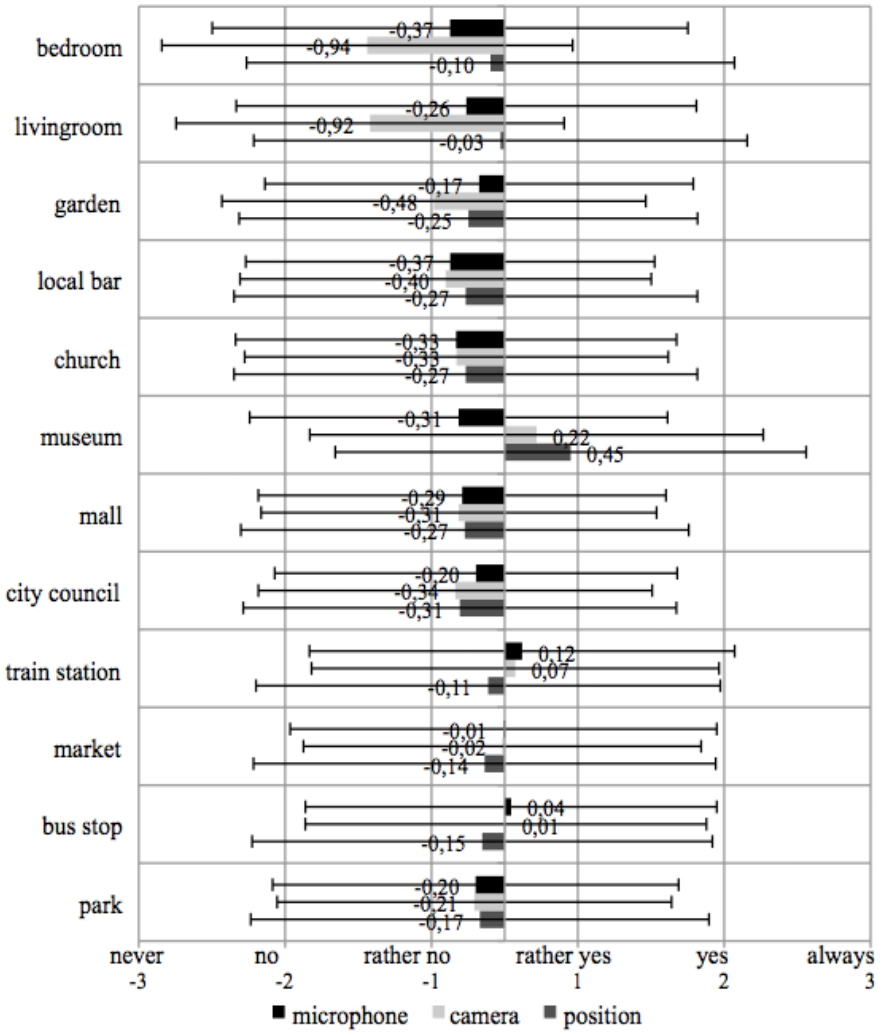


Fig. 6. Mean acceptance with SD by technologies and spaces

4.3 Acceptance of Camera Surveillance and Effects of Health Status

The explorative research approach using correlation analysis indicated the most significant effects on acceptance regarding the camera technology for each space (dependent variables) and the attribute of being healthy or ill (independent variable). Age and gender did not show any significant correlation effects on acceptance.

The acceptance patterns for healthy people can be orientated on the general acceptance results for camera (compare fig.6 with fig. 7)).

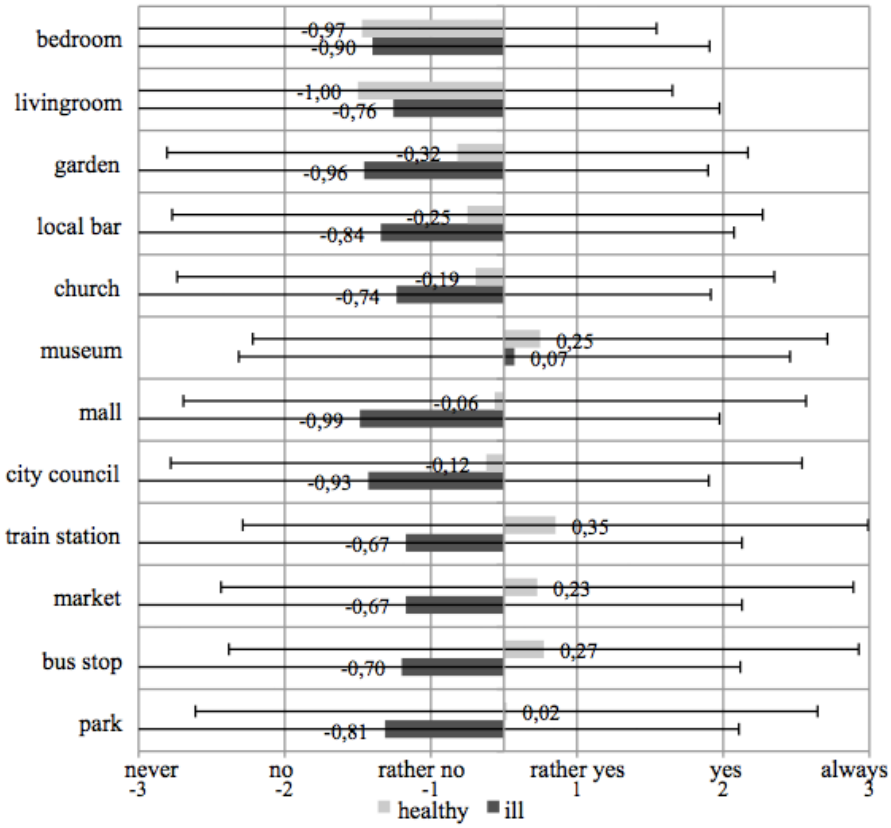


Fig. 7. Mean acceptance with SD of camera surveillance in different spaces, health effect

Ill and frail persons show a significantly lower acceptance for camera surveillance except in private indoor spaces (living- and bedroom, in which both groups do not accept cameras) and the exceptional museum (where both groups would accept cameras, if at all). The camera in garden ($F(1,121)=5.08$; $p<0.05$), local bar ($F(1,121)=8.8$; $p<0.05$) and church (just failed $p<0.05$ significance level) are congruently assessed: No one accepts cameras which is more distinct within ill people. This different acceptance rating grows with changing dimensions (see fig. 1.) from private to public spaces. The distinction in acceptance regarding mall $F(1,120)=21.24$; $p<0.00$) and city council ($F(1,120)=9.72$; $p<0.00$) still increases wherever both are still not accepted. Moving on from indoor to outdoor respective from intimate to more distant spaces, the acceptance not only keeps drifting, observing park $F(1,120)=10.26$; $p<0.00$), market $F(1,121)=20.22$; $p<0.00$), train station $F=(1,121)=15.11$; $p<0.00$) and bus stop $F=1,121)=23.7$; $p<0.00$) – the tendencies even show a turning point for healthy people, who barely could accept these technologies. We therefore can say for the camera as one health monitoring solution and healthy people as regarded user attribute the debated dimensions have a significant influence on acceptance.

5 Discussion and Outlook

In this study we undertook a first exploratory approach regarding the question if and if so persons would accept medical monitoring technologies within urban living environments. The need of understanding individual barriers and perceived benefits of technologies integrated within urban environments is high, especially as increasingly more and older people are living in urban communities, which might be bothered by age-related frailness and which need medical monitoring to maintain mobility and independency at older age.

Summarizing the findings of this exploratory study we see a lot of reluctance to accept medical monitoring in urban environments. Especially the camera is least accepted, followed microphone and positioning systems. Though, acceptance strongly differs depending on the specific location (especially for the camera and the microphone, to a lesser extent for the positioning system). The differentiation between private and public spaces has a strong impact on acceptance, especially for camera surveillance, while the difference between indoor and outdoor spaces plays a minor role.

Referring to well established technology acceptance theories we could replicate the strong effect of technology adaption: surveillance systems for public places (train station, museum), which were first installed ages ago for crime prevention [15], received the highest acceptance. Also, user diversity is a modulating factor of technology acceptance of ubiquitous surveillance technologies. When widening the focus from private living space to public urban quarter, a higher technology acceptance for healthy people was revealed. This might be contra-intuitive on a first sight as healthy persons could see less advantages in health surveillance (technologies) in comparison to ill people, which - according to common preconceptions – are naturally accepting medical technologies, simply because they do not have alternatives [16]. However, the sensitive understanding of illness and age require deeper insights into individual attitudes, wishes, needs, and values. Possibly, ill people could be more sensitive to health monitoring technologies, especially those depending on medical devices, as they might fear stigmatizing and violations of intimacy requirements and data privacy [12]. Future studies will have to find out the reasons of older persons to accept or decline medical surveillance technologies. Still more important we will have to explore possible conditions under which persons would conditionally accept medical technologies implemented within the urban living environment.

Acknowledgements. Thanks to Julia van Heek and Barbara Zaunbrecher for research assistance. This work was funded by the Excellence initiative of German state and federal governments.

References

1. Beard, J.R., Biggs, S., Bloom, D.E., Fried, L.P., Hogan, P., Kalache, A., Olshansky, S.J.: Global Population Ageing: Peril or Promise. World Economic Forum, Geneva (2011)
2. Valverde, J.L.: 2050: A Changing Europe. Demographic Crisis and Baby Friend Policies. Pharmaceuticals Policy and Law 9(1-2), 270 (2007)

3. Moylan, K.C., Binder, E.F.: Falls in OlderAdults: Risk Assessment, Management and Prevention. The American Journal of Medicine 120(6), 493.e1-493.e6 (2006), doi:10.1016/j.amjmed.2006.07.022
4. Gurley, R.J., Lum, N., Sande, M., Lo, B., Katz, M.H.: Persons Found in Their Homes Helpless or Dead. N. Engl. J. Med. 334, 1710–1716 (1996)
5. Osvald, F., Wahl, H.-W.: Dimensions of the Meaning of Home in Later Life. In: Rowles, G.D., Chaudhry, H. (eds.) Home and Identity in Later Life: International Perspectives, pp. 21–45. Springer, New York (2005)
6. Ziefle, M., Himmel, S., Wilkowska, W.: When your living space knows what you do: Acceptance of medical home monitoring by different technologies. In: Holzinger, A., Simonic, K.-M. (eds.) USAB 2011. LNCS, vol. 7058, pp. 607–624. Springer, Heidelberg (2011)
7. Slobogin, C.: Public Privacy: Camera Surveillance of Public Places And The Right to Anonymity. Mississippi Law Journal 72 (2002), doi:10.2139/ssrn.364600
8. Surette, R.: The thinking eye: Pros and cons of second generation CCTV surveillance systems. Policing: An International Journal of Police Strategies & Management 28(1), 152–173 (2005)
9. European Commission (EUR 22030), "Eu Research on Social Sciences and Humanities 2004: On the Threshold to Urban Panopticon? Analysing the Employment of CCTV in European Cities and Assessing its Social and Political Impacts – URBANEYE" Final report. Office for Official Publications of the European Communities, Luxembourg (2007)
10. Alagöz, F., Ziefle, M., Wilkowska, W., Valdez, A.C.: Openness to accept medical technology - A cultural view. In: Holzinger, A., Simonic, K.-M. (eds.) USAB 2011. LNCS, vol. 7058, pp. 151–170. Springer, Heidelberg (2011)
11. Arning, K., Ziefle, M.: Different Perspectives on Technology Acceptance: The Role of Technology Type and Age. In: Holzinger, A., Miesenberger, K. (eds.) USAB 2009. LNCS, vol. 5889, pp. 20–41. Springer, Heidelberg (2009)
12. Wilkowska, W., Ziefle, M.: Privacy and Data Security in E-health: Requirements from Users' Perspective. Health Informatics Journal 18(3) (2012)
13. Beier, G.: Locus of control when interacting with technology (Kontrollüberzeugungen im Umgang mit Technik). Report Psychologie 24, 684–693 (1999)
14. Ziefle, M., Himmel, S., Holzinger, A.: How Usage Context shapes Evaluation and Adoption in different Technologies. In: Rebelo, F., Soares, F. (eds.) 2nd International Conference on Human Factors and Ergonomics in Healthcare 2012,, Advances in Usability Evaluation Part II, pp. 2812–2821. CRC Press (2012)
15. Harris, C., Jones, P., Hillier, D., Turner, D.: CCTV surveillance systems in town and city centre management. Property Management 16(3), 160–165 (1998)
16. Ziefle, M., Schaar, A.K.: Gender differences in acceptance and attitudes towards an invasive medical stent. Electronic Journal of Health Informatics 6(2), e13, 1–18 (2011)

The Impact of Explanation Dialogues on Human-Computer Trust

Florian Nothdurft, Tobias Heinroth, and Wolfgang Minker

Institute of Communications Engineering
University of Ulm, Germany
{florian.nothdurft,tobias.heinroth,wolfgang.minker}@uni-ulm.de

Abstract. Maintaining and enhancing the willingness of a user to interact with a technical system is crucial for human-computer interaction (HCI). Trust has shown to be an important factor influencing the frequency and kind of usage. In this paper we present our work on using explanations to maintain the trust relationship between human and computer. We conducted an experiment on how different goals of explanations influence the bases of human-computer trust. We present the results of the conducted study and outline what this means for the design of future technical systems and in particular for the central dialogue management component controlling the course and content of the HCI.

Keywords: Computer applications, Knowledge based systems, Cooperative systems, Adaptive systems, Expert Systems.

1 Introduction

Advances in human-computer interaction based technology enable the vision of mobile or ubiquitous technical systems accompanying users in their daily life. These systems have the possibilities to serve as a personal assistant due to the potential long-term relationship between a human and a technical system. Personal assistants should have the potential to solve complex problems the user is faced with daily or solely and which require significant interaction. However, this paradigm of interaction requires a working relationship between the human and the technical system. Such a relationship is characterized by a user's cooperativeness during interaction and his trust in the technical system.

Human-computer trust has shown to be a crucial point in keeping the user motivated and cooperative. The users' trust in a technical system will decrease if he does not understand system actions or instructions (Muir, 1992). This may lead to a change in the willingness to interact or in the worst case scenario to an abort in interaction and use (Parasuraman & Riley, 1997). However, providing explanations can help to prevent a decrease of trust (Glass, McGuinness, & Wolverton, 2008).

Similar to explanations in human-human interaction, explanations in human-computer interaction pursue a certain goal. Explanations are given to clarify, change or impart knowledge with the implicit idea to align and adapt the mental models of

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 59–67, 2013.

the participating parties. In this case the users' constructed mental model of the system has to be adapted to the correct mental model (i.e. the designed behavior, knowledge and reasoning) of the system. This means, that not only the knowledge of the user has to be adapted to match the system required knowledge, but the behavior of the system has to be explained to the user in order to keep his transparency in the system. For example, if the users' believed reasoning process of the system does not match the real reasoning process, the occurrence of not understandable or as wrong perceived situations seems inevitable. The correction of the users' mental model can be done by providing explanations.

In order to pursue a particular objective an explanation goal (see Table 1 for a listing of explanation goals) has to be selected. However, a complex explanation can pursue more than one goal at a time. For example, one can think of a learning explanation which contains a conceptual explanation as well. Additionally a simple explanation does not necessarily have to pursue only one goal. The mapping of the presented final explanation to the inherited explanation goal does not have to be a one to one mapping. This means that the effect of a given explanation is not limited to the originally intended goal, but may implicitly pursue other goals as well.

Table 1. The different goals an explanation can pursue

Goal of Explanation	*Description*
Justification	Explain the motives of the answer?
Transparency	How was the systems answer reached?
Relevance	Why is the answer a relevant answer?
Conceptualization	Clarify the meaning of concepts
Learning	Learn something about the domain

For our experiment we concentrated on justification and transparency explanations. Justifications are the most obvious goal an explanation can pursue. The main idea of this goal is to provide support for and increase confidence in given system advices or actions. The goal of transparency is to increase the users understanding in how the system works and reasons. This can help the user to change his perception of the system from a black-box to a system the user can comprehend. By this, the user can build a mental model of the system and its underlying reasoning processes. Selecting the appropriate goal of explanation based on users' human-computer trust is an unprecedented approach because existing studies concentrate on trust as a one-dimensional concept. However, Trust is multi-dimensional and consists of several bases. For human relationships, Mayer (Mayer, Davis, & Schoorman, 1995) defined three levels that build the bases of trust: ability, integrity and benevolence.

For human-computer trust (HCT) Madsen and Gregor (Madsen & Gregor, 2000) constructed a hierarchical model (see Fig. 1). They tried to separate trust into nine basic constructs but eliminated four constructs because of representative or discriminative issues. This results in five basic constructs of trust, with two major components (cognitive- and affect-based components) and expected relationships modeled between them. However, as Mayer already stated, the bases of trust are separable, yet

related to one another. All bases must be perceived high for the trustee to be deemed trustworthy. If any of the bases does not fulfill this requirement, the overall trustworthiness can suffer (Madsen & Gregor, 2000).

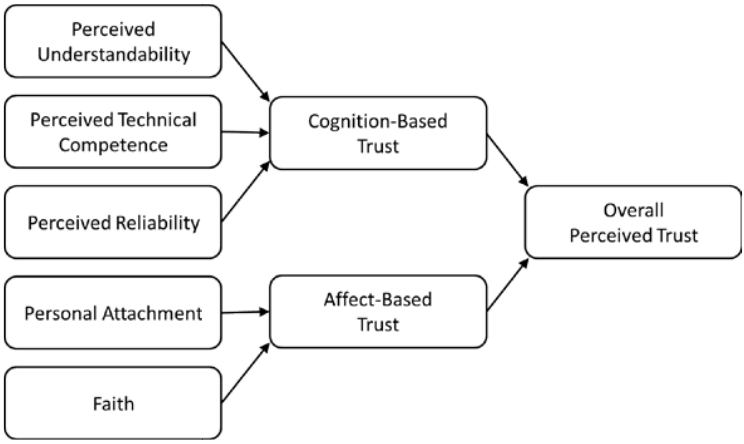


Fig. 1. In this for human-computer trust constructed model, personal attachment and faith build the bases for affect-based trust and perceived understandability, perceived technical competence and perceived reliability for cognition-based trust (Madsen & Gregor, 2000)

If we want to use explanations to influence the human-computer trust relationship in a directed and not arbitrary way, we need to find the most effective mapping of explanation goals to HCT bases (see Fig. 2). This means, that we have to identify which goal of explanation influences which base of trust in the most effective way. Thereby, undirected strategies to handle HCT issues can be changed into directed and well-founded ones, substantiating the choice and goal of explanation.

In our experiment we wanted to test how the different goals of explanation do influence the bases of trust in unexpected, not understandable situations in human-computer interaction (HCI). The main idea was to influence the HCT relationship in a negative way and to analyze how different explanation goals can help to remedy or reduce occurring trust issues to prevent the user from losing the willingness to interact with the system.

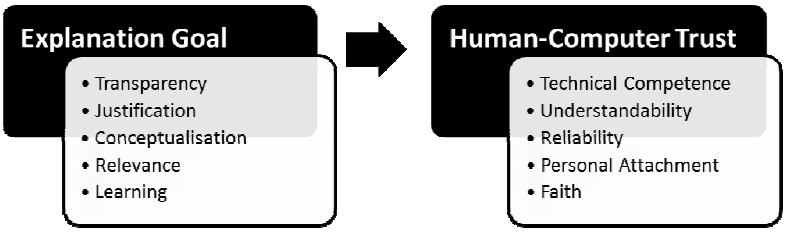


Fig. 2. The main goal is to find the most effective mapping between goals of explanation and the bases of human-computer trust

2 The Experiment

The setting of the experiment was a web-based simulation of a nuclear power plant control room. The subjects had to accomplish several rounds of interaction in which they had to solve pre-defined tasks. During those rounds they were assisted by a virtual anthropomorphic assistant (Lang & Minker, 2012) which helped the user proactively with the upcoming tasks. The user interface represented the controls of the nuclear power plant by distributing control room functionalities over various tabs (see Fig. 3).

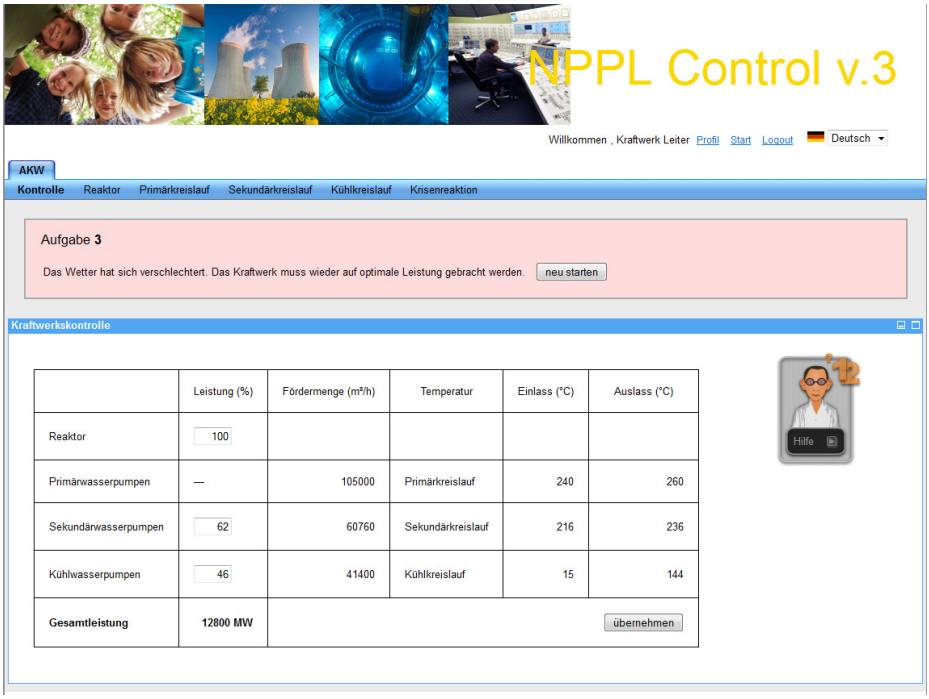


Fig. 3. This is a screenshot of the experiment interaction interface. On the right we can see the agent in sleep mode. Distributed over the tabs are the nuclear power plant functionalities. Beneath is the text describing the task the user has to accomplish. In this case the weather is getting worse and the power plant has to be adjusted to provide more power.

However, during selected rounds, the simulation reacted unexpectedly. The main idea behind this was to provoke a decrease in human-computer trust between the human and the machine. By this we wanted to test how different goals of explanation can help to prevent or reduce the expected trust loss. In total 60 test persons took part in the experiment. For each kind of explanation 20 persons were tested with an evenly distributed number of males and females. However, due to incomplete data only 48 valid subjects remained. The average age was 23.35 with the majority of the participants being students. In order to measure the influences on the bases of

human-computer trust we were using a translated version of the working alliance inventory questionnaire modified by Madsen and Gregor (Madsen & Gregor, 2000) for the measurement of human-computer trust. The original questionnaire (Horvath & Greenberg, 1989) measures which trust and belief a therapist and patient have in each other in achieving a desired outcome. The HCT questionnaire was adapted to our needs and consisted of 15 items (three items for each base of trust).

In total the subjects had to complete seven tasks. For example, the nuclear power plant had to be controlled to output a certain amount of power. However, three of the seven tasks were interrupted by unexpected situations. For example, a water pump was broken or some control rod elements were defect. These situations were meant to be incongruent to the users' mental model of the system and therefore not understandable and unexpected. The course of the experimental design can be seen in Table 2.

Table 2. The Course of the experiment regarding unexpected situations and provided HCT-questionnaires

Round	*1*	*2*	*3*	*4*	*5*	*6*	*7*
Error / unexpected Situation			x		x	x	
HCT Questionnaire		x	x	x	x	x	

In the beginning we wanted the user to accustom to the system. Therefore, the first questionnaire was presented after the second task. In the third, fifth and sixth round the task was interrupted by an unexpected system error (see Fig. 3).

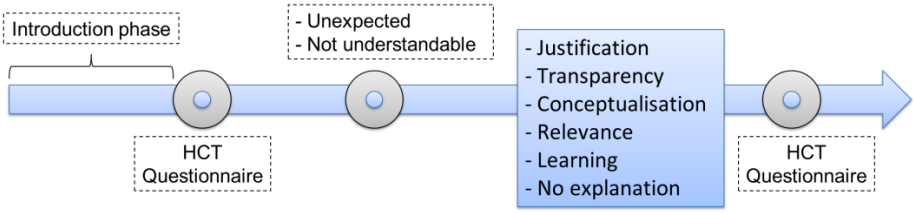


Fig. 4. Rounds with unexpected situations in the experiment: In the first study we examined justification, transparency and no explanation

As mentioned before the system reaction was either augmented by a transparency or justification explanation. The baseline was a group provided with no additional explanation. These experiments allowed us to determine, whether our constructed unexpected situations did influence the HCT negatively. Our hypothesis was that both goals of explanation would perform better in terms of keeping trust than no explanation at all. Additionally, we assumed that justification explanations would help especially the bases of technical competence and understandability. For transparency explanations we expected influences on the bases of understandability and reliability.

3 Results

The first problem we encountered was that the unexpected situations did not induce the anticipated trust loss. In our opinion, this was either due to the too good interaction and help of the system represented by the virtual agent or the introductory phase was too short to build a trust relationship between man and machine.

Observing the data we did not find any significant differences between providing the user with no explanations, justifications or transparency explanations. Especially the development over time seemed rather arbitrary in terms of system errors influencing the human-computer trust relationship. However, when analyzing the data we found some gender differences (see Fig. 4).

Concerning the base of the perceived reliability we found a marginal significant ($H^2 = 2.9, p < .08$) difference when using transparency explanations (4.29 for males to 3.49 for females). When providing justifications we got a significant difference ($H^2 = 4.0, p < .05$) concerning the perceived faith between males (3.2) and females (4.12). When observing only female subjects we could prove a marginal significant ($U = 8.5, p < .09$) difference between providing no explanation (3.9) compared to providing justifications (4.57) for the base of reliability.

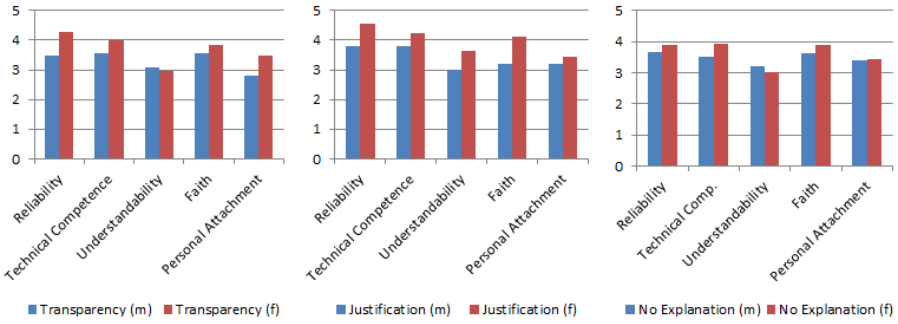


Fig. 5. Results for the different types of intervention divided on gender

Closer examination of the data revealed some further tendencies. However, due to the study design there were not sufficient numbers of subjects to analyze the male to female differences to draw valid conclusions. In the next chapter we will discuss the results and mention some of the rather interesting tendencies which we hope to address in more detail in future work and experiments.

As mentioned before the situations meant to influence HCT negatively did not serve their purpose. Despite experiencing not understandable situations the help provided by the agent was sufficient to handle the occurring problem. The agent provided a step-by-step tutorial on how to overcome the experienced problems. Therefore, in a follow up experiment we want to separate the occurring system error from the task the subject has to accomplish. Additionally, we plan to extend the introductory phase. Every type of task the user has to execute later on, should be done in a comparable way in the beginning. This way we hope to build a more complete users' mental model of the system.

Taking a closer look at the development of perceived reliability and perceived technical competence (see Fig. 5) when providing transparency explanation, we observe that for male subjects the curve has rather negative tendencies. Compared to that, females seem to be influenced only in a marginal way by transparency explanations, at least in our experiment.

Providing justifications seemed to benefit the perceived reliability regardless of gender. For perceived technical competence females seemed to benefit from justifications (as well see Fig. 5). As we lost in total 10 female subjects to incomplete questionnaires or quitting of the experiment (compared to 2 males), the number of females per explanation goal was limited. However, the results give some evidence that the gender aspect of explanation goals is worth more thorough investigation in the context of our follow-up studies.

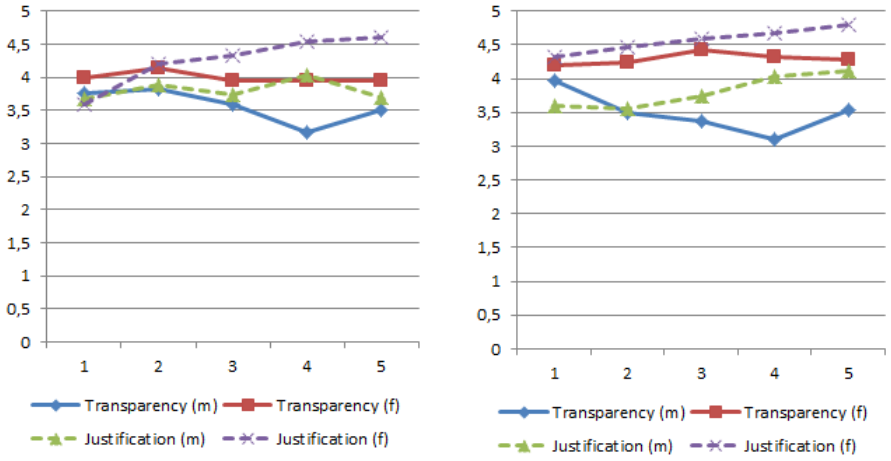


Fig. 6. The values of perceived technical competence on the left and perceived reliability on the right as given by the HCT questionnaire

Despite the setback of not influencing the HCT negatively, we can state that different goals of explanation do influence the bases of trust in a particular way. For example, justifications do influence the perceived reliability of the user towards the system (i.e. Fig. 5). Despite being not significant, this could be a first indicator that goals of explanation influence particular bases of trust.

4 Discussion

Technical systems meant to accompany and assist the user in his daily life are prone to the consequences of users' loss of trust. Personal assistants rely on users' cooperativeness and motivation to interact with the system and share relevant information. This is necessary to facilitate the adaptation and personalization of the technical system to the individual user. Especially future technical systems will have to adapt to a

user's capabilities, preferences, requirements, and current needs. As these continually available assistants should appear competent, co-operative, and reliable to their users the relationship between human and technical system has to be healthy and trustworthy. If technical systems should change from helpers for simple, domain dependent tasks to individualized continually available assistants for complex and wide-spread domain independent tasks the dialogue between human and technical systems can be a typical bottleneck. The accomplishment of complex tasks requiring extensive dialogue depends on the users' willingness and motivation to interact. Maintaining the trustworthiness of the systems is, as mentioned earlier, one of the key factors to keep the user willing to interact with the system. Continually available and assisting systems possess a bunch of relevant information on the user, which can facilitate the preservation of human-computer trust. Gender, education, technical background and history of interaction can help to recognize situations prone to HCT-decrease and to select the probably most suitable intervention to remedy possible negative effects.

As presented in this paper, we found significant results as well as important evidence that the loss of trust has to be handled in a directed, not arbitrary way. Diverse types of explanations can influence the bases of trust in directed way. A well-founded selection of the most effective type of explanation will help to handle occurring trust issues in a non-arbitrary way. Additionally, the experiment indicates that females and males seem to react in a different way to types of provided explanations. As this does not have to be the only additional factor in the explanation selection, future research has to take other factors as education or technical background into account as well. Especially technical background is known to have an effect on explanations. The content, complexity and form of the explanation have to be adapted to the level of technical background and knowledge a user possesses. For example, expert and novice users profit differently from the diverse goals of explanation (Lim, Dey, & Avrahami, 2009). In our opinion, the goals of explanation may influence the bases of trust differently regarding the expertise as well.

However, there are still some missing pieces to be able to use these and upcoming results in a technical system for adaptation and individualization. Knowing how to react in the probably most effective way to a decrease in a base of trust is important, but the information which base of trust was influenced is essential as well. For this we need to categorize the context information in situations of trust loss. If we can state with a certain probability, that a particular context along with unexpected situations influences a specific base or bases of trust, we can react in an appropriate manner. This will enable us to include the results in an existing architecture to handle human-computer trust issues by providing explanations (Nothdurft, Bertrand, Lang, & Minker, 2012).

5 Conclusion and Future Work

In this paper we outlined why it is necessary to consider human-computer trust in technical systems. Especially in systems meant to be personal, individualized and continually available assistants, the long-term relationship between human and technical system is highly relevant. Unexpected or not understandable situations have shown to influence the users' HCT negatively. Research has shown that explanations

can help to remedy negative effects occurring in these situations. However, the mapping between the different goals of explanations and the diverse components of human-computer trust remains unclear. Therefore, we presented an experiment on how different goals of explanation influence particular bases of human-computer trust. We found indication that indeed differences exist in the effects of explanation goals on the bases of trust. Additionally, we did find some gender aspects, which seem to be worth analyzing more extensively in follow-up experiments. Therefore, it seems worthy to consider individual and situational parameters when providing explanations in critical situations.

Acknowledgements. This work was supported by the Transregional Collaborative Research Centre SFB/TRR-62 "Companion-Technology for Cognitive Technical Systems" which is funded by the German Research Foundation (DFG). Additional thanks to M. Attenberger, L. Knecht, N. Maucher, M. Partsch and F. Rueck for conducting the study.

References

Glass, A., McGuinness, D.L., Wolverton, M.: Toward establishing trust in adaptive agents. In: IUI, pp. 227–236. ACM (2008)

Horvath, A., Greenberg, L.: Development and validation of the Working Alliance Inventory. Journal of Counseling Psychology, 223–233 (1989)

Jörg, C., Anders, K.-P.: Designing Explanation Aware Systems: The Quest for Explanation Patterns. In: Explanation-Aware Computing – Papers from the 2007 AAAI Workshop, pp. 20–27. AAAI Press, Vancouver (2007)

Lang, H., Minker, W.: A collaborative web-based help-system. In: Proceedings of the 2nd International Conference on Web Intelligence, Mining and Semantics, pp. 61–65. ACM, Craiova (2012)

Lim, B.Y., Dey, A.K., Avrahami, D.: Why and why not explanations improve the intelligibility of context-aware intelligent systems. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 2119–2128. ACM, Boston (2009)

Madsen, M., Gregor, S.: Measuring human-computer trust. In: 11th Australasian Conference on Information Systems (2000)

Mayer, R.C., Davis, J.H., Schoorman, F.D.: An Integrative Model of Organizational Trust. The Academy of Management Review, 709–734 (1995)

Muir, B.M.: Trust in automation: Part I. Theoretical issues in the study of trust. In: Ergonomics, pp. 1905–1922 (1992)

Nothdurft, F., Bertrand, G., Lang, H., Minker, W.: Adaptive Explanation Architecture for Maintaining Human-Computer Trust. In: 36th Annual IEEE Computer Software and Applications Conference (COMPSAC 2012), pp. 176–184. IEEE, Izmir (2012)

Parasuraman, R., Riley, V.: Humans and Automation: Use, Misuse, Disuse, Abuse. Human Factors. The Journal of the Human Factors and Ergonomics Society, 230–253 (1997)

Sormo, F., Cassens, J.: Explanation goals in case-based reasoning. In: Proceedings of the ECCBR 2004 Workshops, pp. 165–174 (2004)

Trustworthy and Inclusive Identity Management for Applications in Social Media

Till Halbach Røssvoll and Lothar Fritsch

Norwegian Computing Center (Norsk Regnesentral)
Gaustadalléen 23, 0314 Oslo, Norway
{Till.Halbach.Rossvoll,Lothar.Fritsch}@nr.no

Abstract. We describe a prototype for inclusive and secure identity management regarding a bill sharing application in social media. Beginning with the principals of universal design, and involving groups of users with impairments, we designed a set of alternative authentication methods based on OpenID. This work explains the scenario and the particularities of designing a trust, security, and privacy infrastructure with a high degree of usability for diverse user groups, and which is aligned with the requirements from regulatory frameworks. The user trials show that several authentication alternatives in multiple modalities are welcomed by impaired users, but many have restrictions when it comes to payments in the context of social media.

Keywords: Authentication, Authorization, OpenID, Identity Management, Social Media, Payment, Accessibility, Usability, E-Inclusion.

1 Introduction

Online payment applications naturally require a high level of trust by the user. This applies in particular to payment services inside social media, which are, as of today, typically viewed as being insecure and open, in contrast to for instance secure and privacy respecting internet banking [1].

The e-Me project [2] focuses on this trust and aims at providing accessible, multimodal, and adaptive authentication and authorization methods for social media that are usable for all users. In an integrated social-payment application connected to online banking, an OpenID provider has been developed by means of inclusive-identity management methods. The provider is used for both the social-media access control and the embedded payment service.

So far, the trust issue has been discussed in the HCI community broadly with respect to usable privacy and security, risk and online trust, considering different objects of trust, e.g., websites, companies, and individuals [3–7]. Furthermore, trust has been discussed generally as a factor as a part of the user experience, e.g., [8, 9], and as part of trust evaluation strategies [10]. This work places the term trust inside a particular case, a payment application for social media and links it to identity management in terms of authentication and authorization. Parts of this work have been presented at previous occasions: [11] briefly introduces

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 68–77, 2013.

the PayShare application mentioned further below, and [12] extends the topic with more details on the e-Me project and a detailed discussion of security and privacy measures [13] of the PayShare application. The novel contribution of this paper is an in-depth description of the accessible OpenID server and a thorough discussion of trust aspects of the entire solution.

The work is structured as follows: After the problem description and a brief overview of the provided solutions, trust considerations are discussed in detail. Then, best practices for design for trust are presented, before giving a detailed description of the OpenID server and the description of conducted user trials. Finally, the conclusion is drawn.

This work is funded by the Research Council of Norway through the VERDIKT program under contract no. 201554.

2 Problem Analysis

The formulation of the objective and constraints of the prototypes was put forward in the description of the Norwegian research project e-Me [2]. The solutions should be suitable for real-life use, including the context of social media, be applicable to authentication and authorization likewise, they should be accessible and offer a high degree of usability, and they should avoid to compromise privacy, security [13], and to offend legal frameworks. Additional constraints regarding the honoring of universal-design and legal frameworks are summarized in [14] and [15], respectively.

Prior to any development, a number of key challenges was identified that had been pointed out as open challenges in related research work [16]:

1. The majority of users is suffering from having to handle too many user names and passwords for authentication.
2. Poor accessibility and usability compromise security and privacy and hereby trust.
3. The majority of current authentication mechanisms is not accessible to users with impairments.
4. Users have different requirements and preferences for privacy and security in electronic solutions.
5. Users experience multiple authentication processes in case of frequent authorization as cumbersome.
6. Authentication as used in social media can be applied to privacy and security aware applications without a degradation of the level of security or privacy.
7. When using universal design rather than the legal framework or the established traditions of information technology, and in designing for security and privacy, their incorporation becomes substantially different both in terms of legal compliance as well as on the level of software engineering, information security engineering, and privacy by design.
   - Provide security and privacy mechanisms that adapt to various skills;

- Communication, such as privacy policies, instruction, and terms and conditions must be understandable for persons with a variety of backgrounds;
- Interactions should be rather deterministic, intuitive, and memorable;
- The security infrastructure should be open to special peripheral devices such as audio readers, braille terminals, or interaction interfaces.

Many conflicts between the project's goals turned up. The prioritization of usability and inclusion over security methods created tension both on the technical and on the legal level. By using particular flavors of accessible IDM, at the same time security and DP might become weaker for the users of these methods and the system as a whole. Prioritization between advanced security policies and accessibility had to be done as well as the compilation of a portfolio of alternative authentication methods with focus on accessibility, not on equal security levels. As a consequence, the prototype can accommodate various skill levels, but at the price of very distinct levels of robustness of the security measures. In turn, this approach will make the job of risk assessment and security management of an information security management systems, e.g. according to the ISO 27000 family, more challenging.

Answering these challenges, two main services were developed: An OpenID provider, and an application to share bills among friends and to manage joint payments inside social media, named PayShare. Combined, they have the following properties:

1. OpenID cuts down the numbers of service accounts to remember for the user. In addition, the seamless authentication experience based on a persistent, personally adapted authentication channel matching the users' skills through several services reduce complexity for the users.
2. Full-scale accessibility and a high degree of usability increase the user's trust.
3. Improved accessibility by authentication adaptation in terms of several login alternatives: password, pictures, sounds, pattern, personal questions, additional one-time password (two-factor over separate channel/SMS).
4. User defined threshold for the application of more frequent authentications.
5. Validity of a person's authentication for a user defined time span.
6. OpenID as an authentication means to authorize payments in an financial application inside a social medium.

In the subsequent sections, these contributions are discussed in detail.

3 Considerations Regarding User Trust

Trust is strongly influenced by a service's security measures and privacy awareness. Trustworthy online services require a certain standard regarding user identification, authentication, authorization, role management, and information security and privacy [17]. This applies particularly to a financial application with a virtual wallet like the bill sharing service PayShare, which can be treated

like an internet bank. As a precondition for trust creation, the security mechanisms need to add to the total experience and not get into the way, as often felt by users [18]. From our user groups we learned that identity management concepts and especially authentication are the first major hindrance in participation on social media and other services, as nearly all authentication channels potentially exclude certain user groups [15]. We decided therefore to offer a variety of alternative authentication channels.

The portfolio of authentication methods made the design of security and privacy properties more difficult. The major challenge is the variety of mechanisms with distinct security properties and divergent privacy properties. Depending on a user's choice of mechanisms, the overall security, privacy, and trust framework can differ as compared to other users, and other use cases. Mechanisms may also be changed at any time due to a changing life situation or context. Therefore, it can be difficult to establish a risk management system with static risk assessment. On the other hand, when evaluating the total system security, the strength of the weakest authentication channel must be assumed.

Further issues arise from the identity management unification. By using the OpenID provider as a universal mechanism, PayShare establishes itself as a powerful 3rd party with observation capabilities both towards the social media platform and towards the payment system. In addition, an identification chain from the social media to the PayShare application is established, which continues to 'friends' and to an associated bank. Hence, pseudonymous social media participation is no longer possible. Further issues in trustworthiness are found in the underlying authentication mechanisms, where audio-visual information may leak out to other persons, dependent on the usage situation. It appears that usability and inclusion requirements are in strong tension with security, privacy, and data protection regulation.

Concerning e-inclusion aspects, the requirement of universal design implies a high degree of accessibility and usability of all involved parts of the solutions [14]. The e-Me project considered target groups consisting of users with various impairments, and elderly. Acknowledged impairments were cognitive challenges such as dyslexia, dyscalculi, orientation, learning, and memory problems, sensory challenges like sight and hearing reduction, and motor challenges like trembling hands. Elderly users sometimes have a combination of impairments. However, apart from these groups, the solutions were required to be universally designed, i.e., to be be able to use by virtually all persons.

3.1 Design Measures for Trust

The following design recommendation have been developed and were applied in PayShare to increase the user's trust into the service. They are deliberately held as generic as possible to make them applicable for electronic services in general. They compound both security and privacy aspects, as well as usability best practices.

- Require extra authorization before critical actions. Example: An extra check box has to be marked before for instance claim deletion

- Show concise and comprehensive system messages that explain the general context (what the user is about to do), the concrete task at hand, the requirements needed, and the concrete instructions. Ex.: Instructive messages like "XY has sent you a payment claim", "You are about to reject the claim", "You may need an OpenID address", "Fill in all input fields below, then press .." are shown
- Show brief and comprehensive error messages with both concrete and general help information, and directions to a human contact. Ex: "The system couldn't contact your OpenID provider. It could be caused by .. Please check .. in your settings. If you do not succeed after several trials, please contact .. (link to assistance)"
- Offer a dashboard view to ease overview gaining. Ex.: Showing the status quo, history of events, link to Terms, and link to settings
- Offer multiple easy-to-find links to the profile settings
- Offer several easy-to-find links to Terms&Conditions
- Make all user settings non-mandatory. Ex.: It is not required by the user during the service registration to specify the amount threshold for the additional authentication; instead, the most secure default is chosen (here: additional authentication is needed always), and the user can change this setting later on
- Use safe defaults for all user profile settings
- Only expose particular profile settings on demand. Ex.: Link to the setting "Change the authentication threshold" from the claim payment page, in order for the user to get rid of the "An additional authorization is needed" messages
- Offer several easy-to-find possibilities to delete the user/profile Ex.: Links to the deletion of the user account from at least the settings and the Terms
- Make as many user actions as possible reversible. Ex.: Give particular events a short time span where they can be reverted
- Offer multiple possibilities to delete user data. Ex.: It should be possible for a user to delete all own user entered data
- Anonymize all user data that are impossible to delete. Ex.: In case of an interest conflict regarding the deletion of data, undeletable data should be anonymized. Data minimization is advisable.
- Only show information relevant in a specific situation. Ex.: Do not let the user change the settings when nothing has be changed
- Offer an archive with previous events and actions, comparable to a system log
- Offer a multitude of authentication methods. Ex.: Offer authentication that accounts for sensor, motor, and cognition impairments.
- Hold the design of an OpenID server different from the design of the service to illustrate the mechanisms invoked during an OpenID authentication
- Honor accessibility and usability standards. Ex.: As a minimum, follow the HTML, CSS, WCAG 2, and WAI-ARIA recommendations.
- Run risk assessment concerning the actual strength of the authentication mechanisms versus the value-at-risk in the connected applications. Loss of personal information should be treated as one of the risks.

To wrap up, the user's trust can first of all be created by empowering the user / giving the user access. Second, user trust can be increased by adaptation, where a service provider really "sees" the user and tailors the system according to her needs and preferences. Third, trust can be increased by setting the user in control in terms of informing the user about what is going on, letting the user interact with the system (for verification purposes), and by making the service as transparent, predictable, and reliable/credible as possible. Finally, trust can be increased by smart user support, i.e., by helping the user in case of confusion, insecurity, and system failure.

4 Verification of Considerations

As already mentioned, two prototypes were developed to verify the above considerations: The social-media application PayShare (presented and discussed in [12]), and an OpenID provider. The latter service is the key authentication party and developed with a high degree of accessibility and usability. It is used for login into the social media, registration with the payment application, and for the authorization of payments. To cope with a variety of possible user impairments and preferences, it offers authentication by means of six different login methods, as illustrated in Figure 1 on the following page:

- Password memorizing,
- recognition a series of pictures,
- recognition of a series of sounds (see Figure 2 on page 75),
- pattern drawing,
- knowing the answer to a series of personal questions, and
- a PIN code calculator as a smartphone application.

A good password choice is supported by a password strength calculator detailing the use of lower and capital letters, symbols, and digits. The visual choice is made by picking a sequence (here: five) of pictures out of a set of pictures grouped in categories, such as animals, clothes, and food. Upon login, five different sets of images are presented to the user in sequence, who then has to identify the one correct image from each set. The size of the set of elements (pictures) to pick from and the number of elements the user has to chose can be set as a parameter depending on considerations regarding the number of permutations (in case the order matters) or combinations (when order does not matter) required by a specific application. As an additional security mechanism, a user's account is suspended after four bad trials. The same principle applies to the audio choice, though with sounds and music, and to personal question-answer pairs. With the latter alternative, the user can either choose from the preset collection of questions or formulate own questions. The pattern is drawn by mouse or keyboard on an 8x8 array of points, where between five and 34 points have to be marked. Finally, the PIN code calculator presents a two-channel authentication

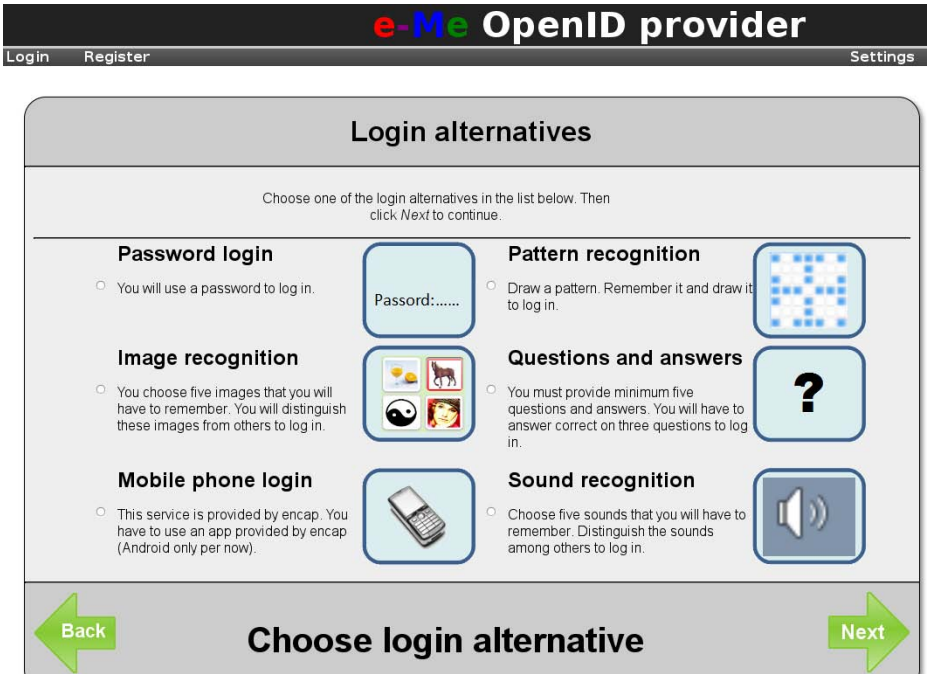


Fig. 1. Screenshot of six login alternatives shown during the OpenID registration

alternative in scenarios where this is required by the service provider (e.g., a bank). It outputs a 6-digit PIN code, which also can be read out load to the user, and which must be typed instead of the user's password. The PIN code is calculated based on a one-time password, which is sent to the phone from the server, and a device identifier.

The user can fill additional data into the OpenID profile, such as birth date, home address, etc. besides the mandatory full name and mail address. More importantly, the prototype also shows that it is possible to personalize the profile by a dedicated settings for the color scheme / contrast, which can easily be extended to other parameters such as font family, font size, and other parameters. Upon authentication, a service may ask OpenID server for all parameters, but it is up to the user to decide which parameters the OpenID server may share with the application. In case a service supports the aforementioned accessibility parameters, the service's design is altered immediately according to the parameter's values when the user returns from the authentication process at the OpenID server. This allows for a "specified once, personalize anywhere" strategy of electronic services. However, in turn, accumulated profile information in OpenID attributes can reveal individual disabilities, and thereby create privacy issues. Disability information might establish processing of medical information, which is regulated strong with respect to data protection in many countries.

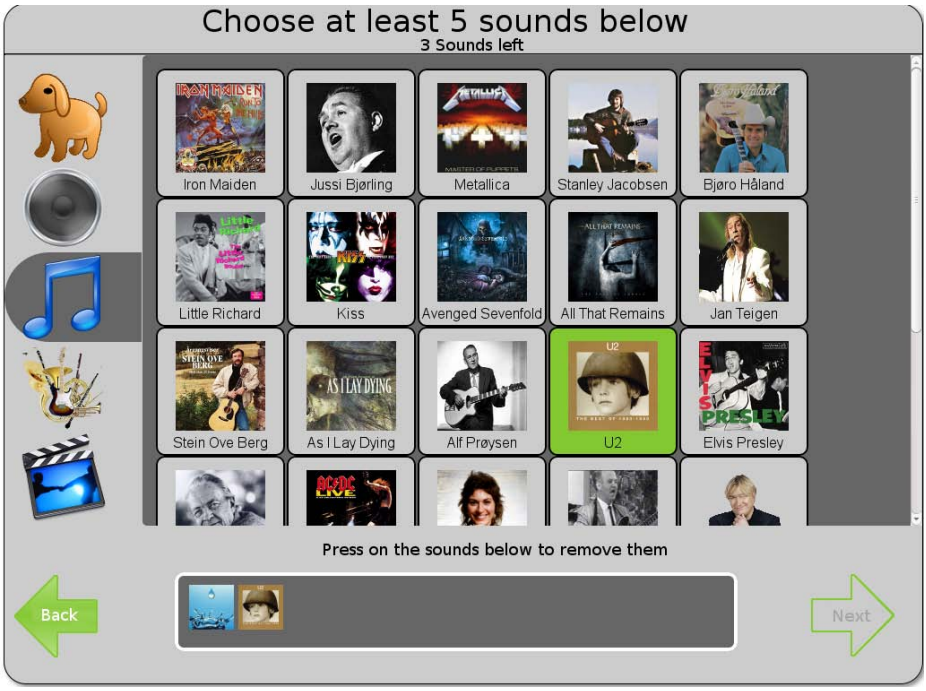


Fig. 2. Screenshot of sound based authentication

4.1 User Trials

In the user trials, eight participants tried out PayShare in combination with the OpenID server. The participants were organized in four groups of two persons and consisted mostly of elderly individuals, some with of them with visual impairments, and some with dyslexia. The groups had first to generate an OpenID account each before it was used for authentication and authorization with the payment sharing application. In the beginning, they were instructed about the situation ("you are NN, have been out eating with XY, you paid the entire bill, and now you claim your money back"), and they were given the names of the other virtual persons. All groups logged in to the given social networking site, which looked very similar to Facebook to provide the proper setting, and after that they were self-driven and only observed by the test leaders. One group filled in a claim, the other groups were notified automatically and had to pay their debt. This "game" quickly lead to new claims forth and back. Once the groups were finished, the participants had to fill out a brief questionnaire.

The first finding is that the deployment of the OpenID server was well accepted, even though none of the participants was familiar with this concept. Only two of the groups, those with sight impaired and dyslectic individuals, chose images as authentication method ("useful", "want that for my token calculator"), the others went for the password alternative ("familiar"). Some of the

sight impaired tried out pattern authentication but soon gave up ("not accessible for us"). It was also commented on the visibility of the pattern and images to others as compared to the concealed password input field. This naturally applies to audio and question-answer pairs as well.

The PayShare service was in general viewed as useful, but the majority had restrictions to make when it comes to trust, mentioning that financial services and social media were in their opinion not compatible. Some of the participants read parts of the ~200-word Terms&Conditions, while nobody had a look at the 300-word privacy text. None of the participants visited their settings, underlining the importance of secure and sensible defaults.

5 Conclusion

We presented a prototype for inclusive and trustworthy authentication and authorization in the context of a bill sharing application in social media. The OpenID server offers a high degree of universal design in terms of six different login mechanisms, each of which based on different modalities. We also discussed how security, privacy, and universal design can increase trust.

A number of key factors is vital to achieve a high degree of trust of the user in the service: Accessibility, adaptation, usability, user control, information availability, interaction, verifiability, transparency, predictability, and reliability/credibility. Accessibility is the most crucial factor as it empowers users in certain situations to use the respective service at all. The other factors increase the feeling of control and thereby the user's trust. The perception of increased trust is not only applicable to users with impairments but rather all, as it is widely recognized that e-inclusion measures for particular focus groups generally increase the service's usability for everybody [19].

References

1. Johnson, M., Egelman, S., Bellovin, S.M.: Facebook and privacy: It's complicated. In: Symposium on Usable Privacy and Security (SOUPS), pp. 24–29 (2012)
2. e-Me Consortium: Inclusive Identity Management in New Social Media, VERDIKT research project no. 201554, Research Council of Norway (2011), http://www.nr.no/pages/dart/project_flyer_e-me
3. Hochheiser, H., Feng, J., Lazar, J.: Challenges in universally usable privacy and security. In: Symposium on Usable Privacy and Security (SOUPS), vol. 2008 (2008)
4. Birge, C.: Enhancing research into usable privacy and security. In: Proceedings of the 27th ACM International Conference on Design of Communication, pp. 221–226. ACM (2009)
5. Dhamija, R., Dusseault, L.: The seven flaws of identity management: Usability and security challenges. IEEE Security & Privacy 6(2), 24–29 (2008)
6. Karat, C., Brodie, C., Karat, J.: Usable privacy and security for personal information management. Communications of the ACM 49(1), 56–57 (2006)
7. Cranor, L.F., Garfinkel, S.: Security and Usablity: Designing secure systems that people can use. Theory in practice. O'Reilly, Sebastopol (2005)

8. Schade, A., Nielsen, J.: Trust and Credibility, 2nd edn. E-Commerce User Experience, vol. 9. Nielsen Norman Group (2000)
9. Corbitt, B., Thanasankit, T., Yi, H.: Trust and e-commerce: a study of consumer perceptions. Electronic Commerce Research and Applications 2(3), 203–215 (2003)
10. Fritsch, L., Groven, A.-K., Schulz, T.: On the Internet of Things, Trust is Relative (chapter 9). In: Wichert, R., Van Laerhoven, K., Gelissen, J. (eds.) AmI 2011. CCIS, vol. 277, pp. 267–273. Springer, Heidelberg (2012)
11. Røssvoll, T.H.: Trust implications for universal design of social-networking applications. In: User-Centered Trust in Interactive Systems Workshop at NordiCHI (2012)
12. Røssvoll, T.H., Fritsch, L.: Reducing the user burden of identity management: A prototype based case study for a social-media payment application. In: Sixth International Conference on Advances in Computer-Human Interactions, ACHI (2013)
13. Fritsch, L.: Social media, e-id and privacy - background for the e-me project. Technical Report DART/02/2011, Norsk Regnesentral (2011)
14. Fuglerud, K.S.: Universal design in ICT services, Trondheim, Norway, 244–267 (2009)
15. Fritsch, L., Fuglerud, K.S., Solheim, I.: Towards inclusive identity management. Identity in the Information Society 3(3), 515–538 (2010)
16. Fuglerud, K.S., Røssvoll, T.H.: Usability and accessibility of personal identification management systems in electronic services. In: Proceedings of eChallenges-2011. IIMC International Information Management Corporation Ltd., Florence (2011)
17. Fritsch, L.: Privacy visualization requirements in the internet of things - a utrustit fp7 ict project note. Technical report, Norsk Regnesentral (Norwegian Computing Center) (2012)
18. Adams, A., Sasse, M.A.: Users are not the enemy: Why users compromize computer security mechanisms and how to take remedial measure. Commun. ACM 42(12), 41–46 (1999)
19. Huber, W., Vitouch, P.: Usability and accessibility on the internet: Effects of accessible web design on usability. In: Miesenberger, K., Klaus, J., Zagler, W.L., Karshmer, A.I. (eds.) ICCHP 2008. LNCS, vol. 5105, pp. 482–489. Springer, Heidelberg (2008)

Towards a Design Guideline of Visual Cryptography on Stereoscopic Displays

Shih-Lung Tsai[1] and Chao-Hua Wen[2]

[1] Chunghwa Picture Tubes LTD., Taoyuan, 334, Taiwan
Tsaihsl@mail.cptt.com.tw
[2] Graduate Institute of Color and Illumination Technology,
National Taiwan University of Science and Technology, Taipei, 10607, Taiwan
chwen@mail.ntust.edu.tw

Abstract. This paper proposed a new visual cryptography scheme with the stereoscopic display which showed and accurately decrypted the hidden information for gray images. Results indicated that contrast ratio and pixel disparity of the decrypted stereo-image were key problems that would impact on the perceived quality of the decrypted image. Next, this research performed a subjective experiment of shifting pixels between both of left and right images to investigate the disparity effects of decrypted information on a full HD stereo-display with film-pattern-retarder technology. In addition, the effects of font size and contrast ratio were addressed as well. Results revealed that the thresholds of pixel disparity were between 2 and 7 pixels. To alphabets, the font size of 50 points was lower boundary to show the decrypted information. To numeric, the font size of 45 points was lower boundary over different contrast ratios.

Keywords: Visual cryptography, Stereoscopic display, Design guideline.

1 Introduction

With the advanced development of science and network technologies, multimedia information is usually delivered over the Internet conveniently. While using secure images, the importance of secure issues cannot be overemphasized because hackers can easy to steal secure information which they need in the Internet. Two popular studies are considered, in general, information hiding is categorized in Watermarking and secure sharing as Visual Cryptography (VC).

Visual Cryptography was proposed firstly by Naor and Shamir [1]. Information of an original image is separated to two or more images. And people just only stacked all encrypted images to decrypt information of original image with Human Visual System. In order to improve the quality of decrypted image, there are many algorithms have been developed to improve the encryption for gray or color images. There are various measures on the performance of many kinds of VC schemes, but rare studies on exact stereo-reproduction in visual cryptography. This study provided a method to

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 78–84, 2013.

amend VC of encryption side and decryption side. Through a series of experiments, visibility requirements of the decrypted information were addressed.

2 Method

Noar and Shamir proposed the (k, n) threshold scheme or k out of n threshold scheme which illustrated a new paradigm in image sharing [1]. In this scheme a secrete image is divided into n share images. With any k of the n shares the secret can be perfectly reconstructed, while even complete knowledge of (k-1) shares reveals no information about the secret image.

This research is expected to answer two questions, 1) what is reasonable disparity in pixels for the decrypted secure image with wearing 3D glasses and 2) what is the limited viewing condition to avert data revealing without wearing 3D glasses. Following sections will describe the basic VC scheme, stereoscopic visual cryptography for gray images and experimental design.

2.1 Basic VC Scheme for Binary Images

The (2, 2) VCS is illustrated to introduce the basic concepts of threshold visual secret sharing schemes as shown in Fig. 1. During the encryption, every secret pixel is split into two shares, and each share belongs to the corresponding share image. For instance, a secret white pixel, the dealer randomly chooses one of the first two rows of Fig. 1 to encode Share1 and Share2. The possibilities of the two encoding cases are equally likely to occur independently whether the original pixel is black or white. Thus, neither Share1 nor Share2 exposes any clue about the binary color of pixel. In the decryption, the two corresponding shares are stacked together, using OR/AND, to recover the secret pixel. Two shares of a white pixel are of the same while those of a black pixel are complementary. Consequently a white pixel is recovered by a share with the stacked result of half white sub-pixels and a black pixel is recovered by all black. Using this basic scheme, the contrast ratio of the decrypted image is reduced results from halving intensity of the white secret pixels.

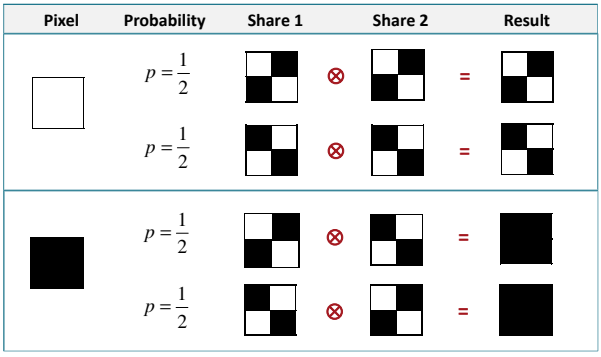


Fig. 1. Encoding scheme for a binary pixel into two shares

2.2 Stereoscopic Visual Cryptography for Gray Images

First we used error diffusion to decompose the gray-scale secret image. And then the halftoning resultant image was processed by the encryption of a basic visual cryptography scheme [2-3]. A 23" LG® D2342P monitor with TriDef® 3D supports was used in this study, the odd rows display the left view and the even rows display the right view in 3D mode. Therefore, it is necessary to rearrange the two encryption share images to match the display format as decrypting the secret images. The VC algorithm for side-by-side file format of stereo polarized displays is described as follows.

Encryption and decryption. Here uses an example of a full HD binary image (1920×1080) to simplify the encryption procedure. First step is to down scale the input binary image. Quarter down-scaling is performed in horizontal direction, e.g., 1920×1080 pixels to 480×1080 pixels. Second is to create the encrypted shares. Based on the basic binary VC, we apply diagonal shares to build a pair of Share1 and Share2. Due to the basic VC scheme, the image size of the Share1 and Share2 are extend to double in horizontal and vertical direction, e.g., 960×2160. After then, we take out the odd row of Share1 and Share2. Third is to rearrange the configuration of Share2. It found that the monitor with TriDef® 3D deals with the odd rows of Share1 to the left eye and the even rows of Share2 to the right eye. For next stage of decryption, it is necessary to switch Share2's the first row to the second row, the third row to the fourth row, and so forth. Fourth is to combine both share images (960×1080) into a full HD encrypted image with side by side format, Share1 is on left side and Share2 is on right side. Fig. 2 illustrates the framework of the stereo visual cryptography scheme for stereoscopic displays in this article. In decryption stage, the encrypted side-by-side image can directly be played on the monitor in 3D mode, and then the secret image can be decrypted clearly without wear glasses. Fig. 3(a)-(c) shows an original secret image, an encrypted side-by-side image and a simulated resultant by our stereoscopic visual cryptography scheme respectively.

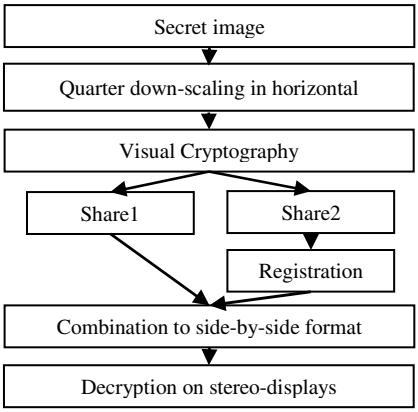
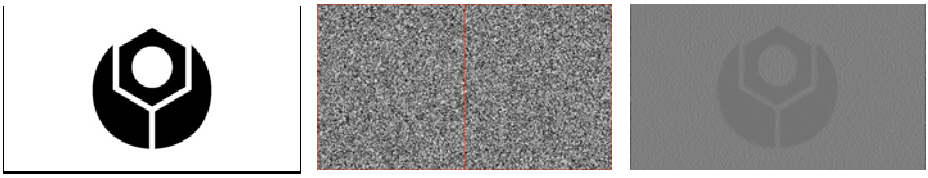


Fig. 2. Framework of the stereo visual cryptography scheme for stereoscopic displays



(a) secret image (b) two encryption shares (Left/Right format) (c) simulated decryption 3D image

Fig. 3. An example of stereoscopic visual cryptography

Changing disparity (pixel shift). For changing disparity, we just move pixels in horizontal of each share with same shift but opposite directions. It is known of increasing disparity can enhance depth perception cues. That is the main reason of decrypted secure image with wear 3D glasses and limited viewing condition to avert data revealing. In this works, it is interesting about how many amount of disparity is a suitable setting for VC. This study investigated six different shift pixels on disparity, such as 0, 4, 8, 16, 24 and 32 pixels. For example, Share1 (left image) is shifted sixteen pixels on the right and Share2 (right image) is shifted sixteen pixels on the left, so the total disparity is thirty-two pixels for side by side image. As the same way, we can finish six disparities for encryption side.

Minimum contrast ratio for binary VC. In this article, we define contrast ratio using Michelson contrast according to the measurements during display profile creation via Eye One Pro by X-Rite®. The Michelson contrast is commonly used for patterns where both bright and dark features are equivalent and take up similar fractions of the area. The Michelson contrast ratio CR is defined as Equation (1),

$$CR = (I_{\max} - I_{\min}) / (I_{\max} + I_{\min}) \tag{1}$$

Here I_{max} and I_{min} representing the highest and dark luminance. The denominator represents twice the average of the luminance. The luminance of white pixel and black pixel were obtained from the color profile. The gamma values of red, green, blue were 2.11, 2.25 and 2.28 respectively. In this article, there are five contrast ratios with gray level are investigated, including 100% (I_{max} = 1; I_{min} = 0), 98% (I_{max} = 0.791; I_{min} = 0.006), 91% (I_{max} = 0.609; I_{min} = 0.028), 73% (I_{max} = 0.452; I_{min} = 0.0691), and 42% (I_{max} = 0.321; I_{min} = 0.13).

2.3 Experimental Design

Participants. Seven male and six female students between 22 and 26 years of age in National Taiwan University of Science and Technology were participated in the experiment. All of them had a normal or corrected to normal visual acuity and passed the stereopsis test according to RANDOT® Stereotests.

Stimuli. Five contents (1920 x 1080 pixels) included three logo icons, a series of Cambria numbers with varied sizes range from 20 to 110 points, and a series of Times New Roman alphabets with varied sizes range from 100 points to 10 points as shown in Fig. 4. At decryption stage, this study investigated six different shift pixels of disparity, such as 0, 4, 8, 16, 24 and 32 pixels. We used Michelson contrast ratio according to the measurements as making the display profile. There were five levels of contrast ratio 100%, 98%, 91%, 73% and 42% respectively. Therefore, there were a total of 150 secret images as visual stimuli in this article.

Experimental task and procedure. The experimental task was a visual detection task. The ambient condition is about average 700 lux. Viewing distance was 50 cm to 70 cm from the display dependent on observer's comfortable viewing position. This study asked participant to watch the contents started from disparity 0 to 32 pixels in five levels of different contrasts from 42% to 100%, then to answer whether perceive the depth clearly or not. At the beginning of the experiment visual acuity and stereopsis were assessed. The experimental procedures are listed as follows.

1. To adjust the monitor set and posture to allow view images at the optimal viewing angle as comfortable as possible;
2. To show a series of a kind of a stimuli on the stereodisplay began to perceive disparity can ascent from 0 pixels to 32 pixels until the subjects do not perceive depth of the original image;
3. To show contents began to perceive disparity can descent 32 pixels to 0 pixels until the subjects can perceive depth of the original image;
4. To record the disparity value, disparity range and calculate the threshold of disparity by font size and contrast ratio;
5. After each stimulus content, a short break could be taken in which the experimenter changed the condition and prepared the experimental software and displays for the next steps. All experimental lasted approximately 90 min.

Fig. 4. Five stimulus contents

3 Results and Discussion

In the following sections, the results about disparity threshold and contrast threshold for 5 contents will be presented, followed by the discussion of relationships between CR and font size.

3.1 Disparity Threshold and Contrast Threshold

Here demonstrates the results of disparity threshold which is the acceptable level or range for applications of stereoscopic visual cryptography. Cormack et al. used a random-dot stereogram and found a cube-root dependency of stereoacuity on contrast at superathreshold levels of contrast [4]. And the difference is probably due to the different spatial-frequency content of the stimuli used in the experiment. Therefore, this study reexamined and also found the disparity threshold was an important comparison among different types of content and contrast.

Fig. 5 depicts the relationship between contrast ratio and disparity threshold. Results revealed that the highest disparity threshold was occurred between CR 91% and 98% and the lowest disparity threshold occurred at CR 42%. Additionally, we confidently concluded that disparity threshold range between 2 and 7 pixels over different contrast ratio. Due to shift two pixels with Share1 and Share2, we cannot see any information on the stereo-display without glasses. Therefore, we need to take out disparity threshold two pixels downward. Fig. 6 illustrates the relationships between different contents and disparity threshold. In evidence, here is the highest disparity threshold in the smile image and the lowest disparity threshold in the Alphabet and Numbers image.

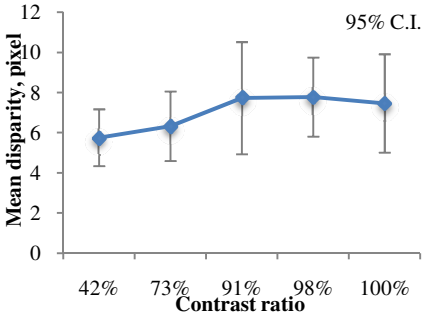


Fig. 5. Plot of the mean disparity threshold as a function of the contrast ratio

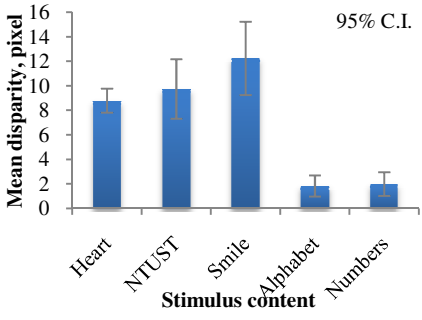


Fig. 6. Mean disparity pixel for 5 stimuli contents

3.2 Relationship between Contrast Ratio and Font Size

Fig. 7 reveals the relationship between font size and contrast ratio in the alphabet size and number character. The alphabet size is blue line with diamond marks; here is the highest font in the 91% contrast ratio and the lowest font in the 100% contrast ratio. The number character is red line with square marks; here is the highest font in the 91% contrast ratio and the lowest font in the 100% contrast ratio.

This experiment tried to find out the disparity threshold of contrast ratio for different content. Results indicated that disparity threshold range is between 2 and 7 pixels over different contrast ratios. For the minimum requirement of alphabet size for

decryption, the font size is 50 points upward (visual angle is about 0.95 degree) in the different contrast ratio. For number characters, the font size is 45 points upward (visual angle is around 0.85 degree) in the different contrast ratio.

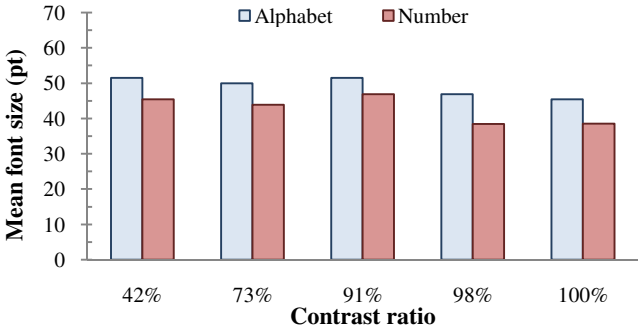


Fig. 7. Mean font size (pt) for different contrast ratio

4 Conclusions

This study proposed a new visual cryptography scheme with the stereo polarized display which can rendering decrypted stereo gray images accurately. For stereo secret images, this method only uses two encryption share images and the decryption can be performed via a side-by-side file format of 3D displays.

Hereby the future works are summarized as follows. First, it is obvious that a lot of time and effort have been dedicated to visual secret sharing using visual cryptography. The trends have been identified within visual cryptography [5]: contrast improvement, share size improvement, wider range of suitable image types, efficiency of VC schemes, ability to share multiple secrets. Finally, because emerging 3D TV, it needs more research on how to use different stereoscopic display with optimal disparity to control suitable and comfortable content protruding.

References

1. Naor, M., Shamir, A.: Visual cryptography. In: De Santis, A. (ed.) EUROCRYPT 1994. LNCS, vol. 950, pp. 1–12. Springer, Heidelberg (1995)
2. Lin, C.C., Tsai, W.H.: Visual cryptography for gray-level images by dithering techniques. Pattern Recognition Letters 24(3), 349–358 (2003)
3. Wen, C.H.: Visual Cryptography on Color Video Displays. In: The 17th International Display Workshop, Fukuoka, Japan, pp. 2105–2108 (2010)
4. Cormack, L.K., Stevenson, S.B., Schor, C.M.: Interocular correlation, luminance contrast and cyclopean processing. Vision Research 31, 2195–2207 (1991)
5. Weir, J., Yan, W.Q.: A comprehensive study of visual cryptography. Springer Transactions on DHMMS 5, 1–10 (2010)

Part II
User Studies

Interface Design for Minimizing Loss of Context in In-situ Remote Robot Control

Jong-gil Ahn and Gerard Jounghyun Kim

Digital Experience Laboratory,
Korea University, Seoul, Korea
{hide989,gjkim}@korea.ac.kr

Abstract. When controlling robot in-situ, the operator's attention is often divided between the scene and the interface. This often causes inefficiency in the control performance. One possible solution to address this attention switch is to employ a camera (or sensor) view (despite being at the site) in which critical parts of the operating environment can be shown side-by-side with the control interface so that the user is not distracted from the either. In addition, when the user switches one's attention away unavoidably and then back to the control interface, the interface can be configured so that user can easily continue the task at hand without momentary the loss of context. In this paper, we describe the design of such an interface and investigate in the possible user attentive behaviors based on it. In particular, we present an experiment that compares three variant forms of interactions: (1) Nominal (no camera view), (2) Fixed (using a camera view and user not allowed to overlook into the scene), and (3) Free (using a camera view but user is free to overlook into the scene). The three approaches represent different balances between information availability, interface accessibility and the amount of attentional shift. Experiment results have shown that all three interaction models exhibited similar task performance even though the Fixed type induced much less attentional shift. However, the users much preferred the Nominal and Free type. Users mostly ignored the camera view, shifting one's attention excessively into the operating scene, due to the lack of visual quality, realistic scale and depth information of the camera view.

Keywords: Human-Robot Interaction, In-Situ Robot Control, Interface Design.

1 Introduction

In-situ remote robot control refers to a situation in which a robot is remotely controlled using an interface device from a near distance (e.g. 2~20m). In-situ remote control often occurs in situations where it is important for the operator to be present at the operating site (e.g. safety reasons, requirement of human expertise, need for effective communication and interaction, slave robots with only limited secondary roles) or when on-board or environment sensors are either limiting or not readily available for complete situation monitoring. Typical examples include bomb diffusing robots, fire-fighting (water gun) robots and robot secretaries.

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 87–96, 2013.

Fig. 1. The three forms of interaction for in-situ robot control which uses the task/scenario based hand-held interface: (1) Nominal (without camera view), (2) Fixed (using a camera view and users not allowed to look over into the operating scene) and (3) Free (using a camera view and but users free to look over into the operating scene).

This is a situation distinct from both close/direct interaction (that e.g. uses voice, touch screen input or gesture) and completely sensor-based tele-operation. Like any interfaces, important factors in the efficient and convenient in-situ remote robot control would include the availability of required information to make the right decision (or situation awareness), access time to the interface for the needed actions and the amount of any distraction or attentional shift between the robot scene and the control interface.

One possible solution to address these factors collectively is to use a hand-held device (such as smart phones and iPad like devices) and at the same time provide a camera/sensor view (from the hand-held device) in which critical parts of the operating environment can be shown side-by-side with the control interface. This way, the user can minimize the times they have to look away (because the camera/sensor view supplies the needed information) from the interface and see directly into the scene, and thus react faster to make proper actions. The user may still be allowed to look over and make direct visual confirmation, if necessary. Thus a provision is still needed to minimize inefficiencies arising from this occasional switching of attention and context.

Note that robots are becoming more versatile and may even require multitasking, and as such also more difficult to control. In addition, hand-held media devices have emerged as attractive and flexible platforms for remote control for their wide popularity and ever-increasing computational and multimedia capabilities [4, 16].

Recently, there has been great amount of research in the field of Human Robot Interaction (HRI), however, concentrated mostly in either direct interaction methods or for tele-operated robots. On the other hand, relatively little research has been focused on remote control interface design for on-site robots. Thus, there is not a clear guideline as how to properly integrate (or if one should), e.g. the camera view and organize the hand-held interface for the best performance and highest usability.

In this paper, we have designed an interface for hand-held in-situ remote robot control that is organized around the task sequence and reconfigured continuously accordingly so that user can be aided in maintaining the task context, and that provides a camera/sensor view to minimize the attentional switch in the first place. Furthermore, based on the interface, we compare three forms of interaction to study user attentive

behavior. The three compared interactions represent different balances between information availability, interface accessibility and the amount of attentional shift.

This paper is organized follows. We first shortly review other related research. Then, we describe the task-oriented reconfigurable hand-held robot interface with the camera/sensor view. Section 4 will describe the details and report the results from the user behavior experiment. Finally, we will conclude the paper with a discussion of the results, a summary and directions for future work.

2 Related Works

2.1 Human-Robot Interaction

With the proliferation of robots, HRI has generated lots of interests among researchers lately. Several guidelines and evaluation metrics have emerged for HRI from the HCI and robotics communities over the years [1, 3, 6, 9, 13, 17]. Most guidelines are quite abstract and general, and also seem to be oriented toward "direct and close" interaction [2, 8, 11] or tele-operation control [10, 12, 14]. Despite relatively less number of studies devoted to in-situ robot control and interaction, many such principles do apply equally. One such guideline is to provide good awareness and less cognitive load [3]. More specifically, Clarkson et al. [1] maintained that natural cues may not be necessary and clear provision of relevant interactive information was more important.

On the other hand, the most typical form of an "indirect" communication with a robot would be through the use of a remote control. Again, we make a note that the indirect approach has been used mostly for tele-robotics or multi-robot control rather than for in-situ bystander driven control. Therefore, the ordinary remote control tends to be small/compact, one way, function-oriented and difficult to use (e.g. button/code/mode based) [5, 7].

2.2 Human-Robot Interface

Based on these HRI principles, guidelines and other requirements, many different forms of human-robot interfaces have been developed [3, 4, 5, 20, 23]. In particular, interface design becomes even more important for efficient control of or interaction with multi-robots or robots with multi-tasks (e.g. navigation, manipulation, interaction, etc.) [19, 21, 25]. In this context, traditional interface devices such as joysticks, joypads, wheels, switches and buttons are being replaced with a new breed that offers rich visual feedback and mobility (e.g. smart phones, tablet PC [26]), gestural input (e.g. Nintendo Wii [22]), and tangibility to compensate for the operator's limited working memory and induce more natural and easier interaction. Among the aforementioned features, visual feedback (usually available through camera views) is one of the most important factors in situation awareness and efficient tele-operation [24]. Zhu et al. studied the tele-operation and multi-tasking performance with regards to different camera views and view control methods [19]. Nielson and Goodrich compared navigation performance between when using camera and map views [23].

3 Interface Design

The robot to be controlled in-situ was originally designed for operator led guidance tasks (e.g. at museums). It was based on a wheeled mobile robot (Adept Pioneer 3-DX) and mounted with a computer, pan-tilt camera, speaker, ultrasonic sensors, and a panning projector (Figure 2). We envisioned that the robot would be operated by a human guide for series of sequential/concurrent tasks involving individual subtasks such as aiming the projector to the wall, selecting and projecting information, controlling the camera to find landmarks (for spatial registration and localization), playing/stopping narrations, positioning the robot, and navigating.

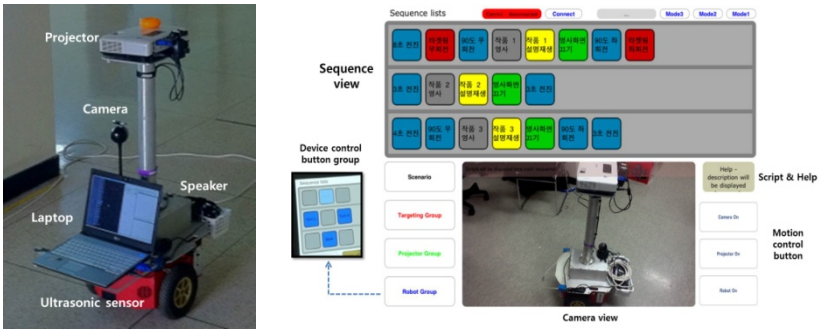


Fig. 2. The museum guidance robot (left) and the scenario based interface (right). The sequence of tasks appear as buttons for the operator to simply activate (under normal circumstances). The tasks are logically grouped. The Camera/Sensor view is in the lower middle (camera remained blank for the Nominal condition).

Since there were relatively many tasks/subtasks to be managed and accomplished as quickly as possible and in order to aid the user quickly regain one's operating context between the occasional view switches, we developed a "scenario" flow based interface. It lays out, in the "Task Sequence View" (top part of the display in Figure 2), pre-configured buttons corresponding to a sequence of tasks to be accomplished for the operator to simply activate. For example, in Figure 2, the eight buttons in the top horizontal line prescribe "Move (robot) forward 1m," "Rotate (robot) 90 degrees to the right," "Turn on projector," "Play narration," and so on. The layout is configured by interpreting an XML based task specification file. The buttons were color-coded or highlighted to e.g. indicate their status (currently pending, already finished) and types (navigation, camera control, projector control, etc.). Therefore, if the robot was to assist a human guide e.g. in a corridor with a number of museum artifacts at known locations, a complete and typical operational scenario can be prescribed and reflected into the interface automatically. Normally, the user would simply use these pre-configured scenario-based buttons to accomplish the given task. In addition the task sequence view interface was used as a way to remind the user what to do without having to memorize the order and specifics of the experimental tasks.

However, note that it was not possible to automatically infer whether a certain task, despite being activated by the operator, really was carried out successfully or satisfactorily. For example, even if a "Move forward 1m" command was issued and executed, the robot may fail to accomplish it due mechanical slippage. In fact, there may be situations that deviate from the normal scenario flow (e.g. new obstacle to avoid). In such a case, the user would need a way to go around the scenario based interface and directly access the various task control interfaces (located in the lower left parts of the display).

The Camera/Sensor view (lower middle part of the display) simply showed the video imagery captured by the camera located in the back of the interface facing toward the operating environment. Note that the camera view was not used for the Nominal interaction model (switched off and remained blank).

The interface was implemented on an Apple iPad [18] which communicated the user commands (with screen touch to various control buttons) to the on-board computer which in turn controlled the robot (movement), projector (on/off, panning), camera (pan and tilt), and voice output. In informal trials, the scenario based interface vastly improved the performance and user satisfaction over the naively designed "bare" interface (see Figure 3). No formal experiment was further conducted as this result was quite expected and obvious.

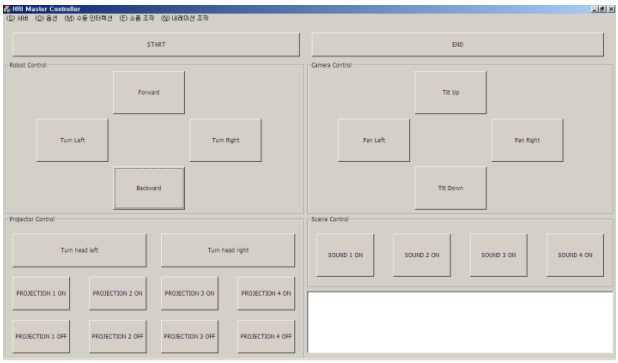


Fig. 3. The naively designed original "bare" interface with simple control buttons

4 Experiment

4.1 Experimental Design and Three Interaction Models

The objective of the experiment was to study the user behavior model in regards to one's shifting of attention between the interface and the actual scene despite the provision of scene view through the camera. Thus we tested three representative interaction models (also see Figure 1) in terms of how the camera view may be utilized (or not).

- "Nominal" type: The Camera view is not provided and the user inspects the operating scene directly.
- "Fixed" type: The hand-held device always aimed toward the target, and the corresponding Camera view is provided. The user is basically not allowed to look over into the operating scene unless absolutely necessary. This represents the usual desktop based tele-operation.
- "Free" type: The user is free to either aim the hand-held device into the scene and use the Camera view or directly look into the operating scene.

The experiment was designed as a 3 x 1 repeated measure within subject, the lone factor being the type of the interface employed. A series of robot control task (details explained in the next section) was given to a user to be carried out by using the given interaction model. The task performance (task completion time and number of clicks made to the interface) was measured and a usability survey was filled out, with a particular focus on the matter of attentional shift. More procedural details are given in the subsequent section.

Table 1. An example sequence of 3 composite tasks to be completed in an experimental session and the experimental space for the user to carry out the sequences of tasks

Sequence	Composite Task Description
1	Move forward to near Placemark 1
	Aim the camera to a marker on the wall
	Turn on the projector / narration (wait)
	Turn off the projector
2	Move forward to Placemark 2 and avoid the obstacle in the way
	Turn on the projector / narration (wait)
	Turn off the projector
3	Move forward to Placemark 3
	Rotate the robot 45 degrees to the right
	Rotate the projector 45 degrees to the left
	Turn on the projector / narration
	Move forward to finish line

4.2 Task

For a given interaction model and the associated interface (given in a balanced order), a series of 3 composite tasks was asked to be completed as fast as possible. An example sequence of the tasks is shown in Table 1 to be carried out in the experimental space shown in Figure 4. To avoid any learning effects as much as possible, the tasks were differentiated between different sessions with a similar degree of difficulty.

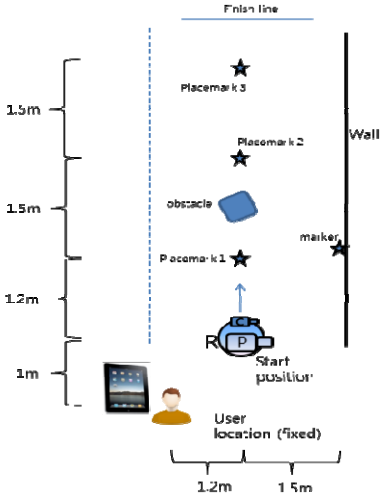


Fig. 4. The experimental space for the user to carry out the sequences of tasks in Table 1

4.3 Detailed Procedure

Eight paid subjects (all men) participated in the experiment with the mean age of 29.5. After collecting one's basic background information, the subject was briefed about the purpose of the experiment and instructions for the experimental task. A short training was given for the subject to get familiarized to the interface and the tasks to be completed.

Each subject went through 9 sessions (3 interfaces types x 3 repeat) lasting for about one hour. Upon starting the experimental session, the subject was asked to use the interface from a fixed position to carry out the task as specified and guided by the Task Sequence View interface. A particular interface type (Nominal, Fixed, Free) was presented to each subject in a balanced fashion.

The task complete time and number of clicks were measured. In addition, the number of times of attentional shift was counted manually (e.g. looking or glancing away from the interface into the operating environment or vice versa). After the subject finished the three sessions, they were asked to fill out the aforementioned survey.

4.4 Results and Discussion

Figures 5, 6 show the task performance, usability (ease of use and fatigue) and the amounts of shifts in the line of sight. Overall there were no statistically significant differences in the task completion nor in the number of clicks made, even though the user made much less (with statistically significant difference) attentional shifts with the Fixed type. There was not any significant difference in the number of shifts between the Nominal and Free type. Most users treated the Nominal and Free type to be mostly equivalent.

We can deduce that the visual information from the direct look into the scene contributed much more to completing the task (which were mostly spatial in nature) in an efficient manner (even when done excessively). Subjects reported, when using the Fixed type, the need to look over to gain more insight into the situation such as proper depth and scale information (which could not be conveyed well with a small camera view). Note that the users of the Fixed type was driven by the experiment to minimize attentional shift, yet it did not improve the task performance measures. It is interesting, despite what the users felt or preferred, similar task performance was achieved with the Fixed type. This was most apparent when the users were trying to avoid an obstacle with the robot.

Attentional shifts occurred much less during relatively stable/simple task operations (e.g. while the robot was moving with sufficient clear room, turning on/off the projector, turning on the narration). Subjects also had a great tendency to make visual confirmation of the task completion (by directly looking into the operating environment) and whether one followed the right task in the given sequence (by looking back into the interface). This generally caused excessive shifting of the line of sight. Usability-wise, subjects mostly grouped the Nominal and Free type together as being more comfortable, easier to use and less tiring than the Fixed type. This, in summary, the results indirectly indicate that the provision of the "high fidelity global" view is very important. For the Free type, the overhead in attentional shifts was compensated by the benefit of the global view for control performance, and reversely for the Fixed type.

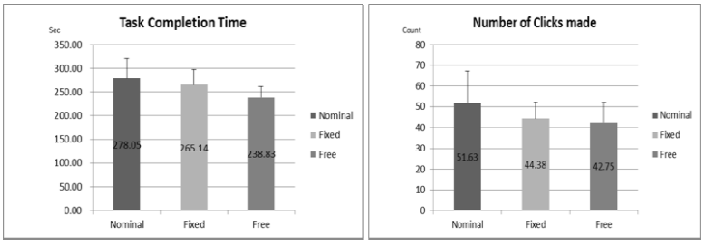


Fig. 5. Task performance results: task completion time (left) and no. of clicks made (right)

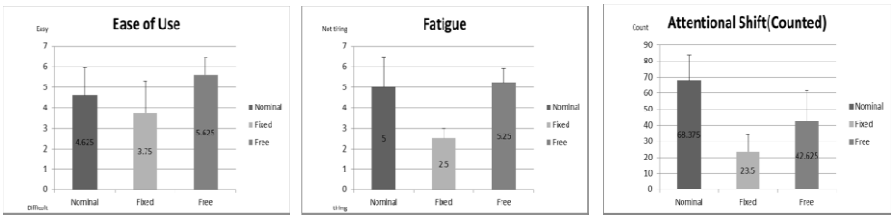


Fig. 6. Subjective ratings(ease of use (left) and fatigue (middle)) and number of attentional shifts(right)

5 Conclusion

In this paper, we have presented a hand-held interface design considering the frequent attentional shift and concern for loss of operational context. We also ran an experiment to investigate the user behavior with regards to the use of camera/sensor view. Our expectation that the attentional shift would play an important and significant role in the efficient and fast task performance was rejected. Despite the much less number of visual shifts, somewhat forced with the Fixed type model, the overall performance was not different from the others. Users much preferred the Nominal and Free type and mostly ignored the camera view despite having to shift one's attention excessively, due to the lack of visual quality, realistic scale and depth information (which were deemed essential in carrying out the mostly spatial robotic tasks). The study illustrates the need to provide a high fidelity global view into the scene. For instance, in a remote tele-operation situation, if a global view can be somehow provided (e.g. a third-person point of view camera), the control performance can be improved. We believe this result can serve as an important and useful guideline for designing control interfaces not only for in-situ robot control, but also for augmented reality and telepresence systems. It reaffirms the importance of scale, depth perception and realism in robotic control.

Despite the ineffectiveness of the camera view in this work, we posit that camera/sensor view that provides information that is not obtainable by the naked eyes will still be important. Different results might be obtained if the camera view had better quality e.g. higher resolution, stereoscopic rendering and larger size. Note that this study focused on the user's choice when both camera and direct imagery were available. Further experiments are needed to investigate the effects toward carrying out non-spatial tasks or toward the quality of the task accomplishment.

Acknowledgement. This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MEST)(No 2012-0009232).

References

1. Clarkson, E., Arkin, R.: Applying Heuristic Evaluation to Human-Robot Interaction Systems. In: Proc. of FLAIRS (2007)
2. Drury, J., Scholtz, J., Yanco, H.: Awareness in Human-Robot Interactions. In: IEEE International Conference on Systems, Man and Cybernetics, Washington, DC (October 2003)
3. Drury, J., Hestand, D., Yanco, H., Scholtz, J.: Design Guidelines for Improved Human-Robot Interaction. In: Proc. of ACM CHI (2004)
4. Muller, F.: iPhone-Robot Command Interface: A Case Study. Semester-Thesis, Swiss Federal Inst. of Technology Zurich (2009)
5. Fong, T., Thorpe, C., Glass, B.: PdaDriver: A Handheld System for Remote Driving. In: IEEE International Conference on Advanced Robotics (2003)
6. Goodrich, M., Olsen, D.: Seven Principles of Efficient Human Robot Interaction. In: IEEE International Conference on Systems, Man, and Cybernetics, pp. 3943–3948 (2003)

7. Hruschak, P.: DIY: Make a Roomba Cell Phone Remote Control (2010), http://www.gadgetell.com/2006/09/diy-make-a-roomba-cell-phone-remote-control/
8. Nourbakhsh, I.: Course Notes for Principles of Human Robot Interaction. Unpublished Work (2007)
9. Olsen Jr., D., Goodrich, M.: Metrics for Evaluating Human-Robot Interactions. In: Proc. of PERMIS (2003)
10. Paulos, E., Canny, J.: Ubiquitous Tele-embodiment: Applications and Implications. International Journal of Human-Computer Studies 46(6), 861–877 (1997)
11. Quigley, M., Goodrich, M.A., Beard, R.W.: Semi-autonomous Human-UAV Interfaces for Fixed-wing mini-UAVs. In: Proc. of IROS (2004)
12. Scholtz, J., Bahrami, S.: Human-Robot Interaction: Development of an Evaluation Methodology for the Bystander Role of Interaction. In: Proc. of IEEE Intl Conf. on Systems, Man & Cybernetics, pp. 3212–3217 (2003)
13. Scholtz, J., Young, J., Drury, J., Yanco, H.: Evaluation of Human-Robot Interaction Awareness in Search and Rescue. In: Proc. of IEEE Intl. Conf. on Robotics and Automation, pp. 2327–2332 (2004)
14. Steinfeld, A., Fong, T., Kaber, D., Lewis, M., Scholtz, J., Shultz, A.: Common Metrics for Human-Robot Interaction. In: Proc. IEEE International Conference on Intelligent Robots and Systems (2004)
15. Networkworld, CES: Battling AR helicopter controlled by iPhone (2012), http://www.youtube.com/watch?v=TSv2ca-IECc
16. Gates, B.: A Robot in Every Home. Scientific American (January 2007)
17. Apple iPad 2 (2012), http://www.apple.com/ipad/specs
18. Zhu, D., Gedeon, T., Taylor, K.: Exploring Camera Viewpoint Control Models for a Multi-Tasking Setting in Teleoperation. In: Proc. CHI 2011, pp. 53–62 (2011)
19. Kawamura, K., Peter, A., Johnson, C., Nilas, P., Thongchai, S.: Supervisory Control of Mobile Robots using Sensory EgoSphere. In: Proc. ICRA 2001, pp. 523–529 (2001)
20. Grandall, J.W., Goodrich, M.A., Olsen Jr., D.R., Nielsen, C.W.: Validating Human-Robot Interaction Schemes in Multitasking Environment. IEEE Transaction on Systems, Man and Cybernetics, Part A: Systems and Humans 35(4) (July 2005)
21. Guo, C., Sharlin, E.: Exploring the Use of Tangible User Interfaces for Human-Robot Interaction: A Comparative Study. In: Proc. SIGCHI (2008)
22. Nielson, C., Goodrich, M.: Comparing the Usefulness of Video and Map Information in Navigation Tasks. In: Proc. ACM HRI, pp. 95–101 (2006)
23. Baker, M., Keyes, B., Yanco, H.: Improved Interfaces for Human-Robot Interaction in Urban Search and Rescue. In: Proc. IEEE International Conference on Systems, Man and Cybernetics, pp. 2960–2965 (2004)
24. Chen, Y.C., Haas, E.C., Barnes, M.J.: Human Performance Issues and User Interface Design for Teleoperated Robots. IEEE Trans. on SMC-Part C: Applications and Reviews 37(2), 1231–1245 (2004)
25. Liang, C.K., Foo, C.H., Lee, Y.S.: Interactive Methods of Tele-operating a Single Unmanned Ground Vehicle on a Small Screen Interface. In: Proc. ACM HRI, pp. 121–122 (2011)

Assessing the Impact of Latency and Jitter on the Perceived Quality of Call of Duty Modern Warfare 2

Rahul Amin[1], France Jackson[1], Juan E. Gilbert[1], Jim Martin[1], and Terry Shaw[2]

[1] Clemson University, Clemson, SC, USA
[2] Cable Television Laboratories, Inc., Louisville, CO, USA
{ramin,fljacks,juan,jmarty}@clemson.edu, t.shaw@cablelabs.com

Abstract. Jane McGonigal stated in her 2010 TED Talk that humans spend 3 billion hours a week playing video games around the planet. Americans alone devote 183 million hours per week to gaming. With numbers like these, it's no wonder why end user demands for bandwidth have increased exponentially and the potential for network congestion is always present. We conduct a user study that focuses on the question: "How much network impairment is acceptable before users are dissatisfied?" In particular, the main objective of our study is to measure a gamer's perceived Quality of Experience (QoE) for a real-time first person shooter (FPS) online game **Call of Duty Modern Warfare 2** in presence of varied levels of network congestion. We develop a Mean Opinion Score (MOS) metric to determine each gamers' QoE. We investigate the following hypothesis: The gamers' perceived QoE correlates to their skill level.

Keywords: Online Gaming, First Person Shooter Games, Network Impairment, Quality of Experience.

1 Introduction

Resource allocation is the process by which network elements try to meet competing demands that applications have for network resources. While broadband cable networks have steadily increased the size of the 'pipe' available to subscribers, end user demand for bandwidth continues to grow at an exponential rate. Therefore, network operators must continue to expand the capacity of their networks in order to avoid congestion. However, it is impossible to avoid occasional network congestion. This congestion is most likely to occur at prime time when most subscribers are active and many different type of applications are simultaneously in use (with many subscribing households using multiple applications simultaneously). The response by each application to congestion differs widely. File transfer applications such as File Transfer Protocol (FTP) or Peer-to-Peer (P2P) file sharing are typically the most flexible in performance requirements. Applications that involve real-time end user interaction, such as video conferencing, VoIP (voice over IP), or online games are less flexible as they generally require low latency. Emerging web-based video streaming applications are generally flexible; however, once performance thresholds are surpassed, perceived

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 97–106, 2013.

quality drops off quickly. Finally, over-the-top HD (high definition) voice and high quality video broadcasts (e.g. live sportscasts) are the most sensitive to fluctuation in bandwidth, latency and loss.

The research described in this paper focuses on network gaming applications. In particular, we study real-time online games. The academic community has studied the impact of latency and loss on real-time games such as first person shooter (FPS) or action games [1]-[7]. The wide range of previous results confirms that mapping network performance metrics such as latency and loss to an assessment of perceived quality is extremely difficult. The challenge is in part caused by the fact that each game responds to latency and loss is unique ways. Further, for a given game, specific modes of operations and even game play scenarios (which change quite frequently) will react differently to latency and loss. Perhaps the most problematic aspect of assessing user's quality of experience is due to the fact that a subjective assessment is a measure of how the gaming experience meets (or fails to meet) a user's expectation. Accurately quantifying expectations is very difficult as a gamer's expectation depends on past history and personal factors.

Our research is motivated by the observation that much of the prior work described in the literature has focused on objective studies. While subjective studies have been considered, there have been limited results. This is primarily because a widely accepted utility function that maps objective metric results to a perceived quality assessment has not been established. While the research described in this paper is primarily exploratory in nature, we consider the following hypothesis: *a gamer's perceived quality of experience correlates to his/her skill level.*

We focus on the widely played FPS game Call of Duty Modern Warfare II. We engage the student population at Clemson University to participate in a large scale experimental study designed to explore the impacts of network impairments on the perceived quality. We collect information from the participants including their self-assessed gaming expertise, scores achieved during the testing session, and assessment of perceived quality. Our results do support our hypothesis by clearly showing that user's level of experience defines and determines the sensitivity to latency. However, we find 'expert' gamers who are presumably the most sensitive to latency are also the most adept at compensating for impaired conditions.

This paper is organized as follows. Section 2 presents a literature survey and summarizes relevant related work. Section 3 provides a description of our experimental methodology. Section 4 documents our results and provides an analysis of our results. Finally, Section 5 provides conclusions including a discussion of the limitations of our study and possible next steps.

2 Background

The broad genre of multi-player, online games includes non-real-time games such as board games or fantasy games as well as real-time games such as first person shooter,

sports, or real-time strategy. The work in [5] presents a useful taxonomy for classifying real-time online games. The authors correlate the game's relationship between latency requirements and 'precision' (either movement or shooting) requirements. Further attributes include the interaction model and the game perspective. The interaction defines how a player interacts with the game world. The authors indicate the interaction is either through an avatar or with an omnipresent model. The avatar model exists at a particular location in the game context and is only involved with local interaction. The omnipresent model provides players a global view of the game. The game perspective is tied to the interaction model as it defines how players view the game world. For avatar interaction, such as in shooter games, a first or third person shooter perspective is common. For omnipresent interactions, such as in role-playing games, the perspective might be variable. A player might have access to aerial views as well as first or third person views.

Because real-time online games are sensitive to network and system performance, there has been a significant amount of prior academic research in this area. Many of these studies share our goals and methods. Works such as [1]-[7] all seek to understand the impact that loss and latency have on real-time online games. A commonly used evaluation approach is to emulate loss and latency in a controlled test-bed environment. Either objective measures such as game results or subjective tests based on Mean Opinion Scores (MOS) are used. The impacts of loss are arguably not as significant as those caused by latency because most games are designed to tolerate some amount of loss either by sending game state or control commands frequently or by utilizing error correction. Therefore, much of the research presented in the literature has focused on the impacts of latency.

In [1], the authors found that the ability to carry out 'precise' movement and shooting tasks is not significantly affected by loss rates or latency that are typical for the FPS game Unreal. They did observe players scores (in terms of kills and deaths) in some game modes drop as the latency increased beyond 100 milliseconds (ms). In [2], the authors focus on the effects of latency on World of Warcraft. They find that because the game is real time strategy (rather than first person shooter), a high level of induced latency (well beyond 100 ms) is noticeable but has negligible effect on the outcome. Other works have used methods involving subjective tests where gamers were asked about their perceived quality. In [7], as round trip times exceed 150 ms, the performance of Halo gamers decreased by about 50% while the perceived quality was still high. Further, games with over 500 ms of delay and 4% loss were considered unplayable. The work in [3] shows that in some scenarios (i.e., play situations of a specific FPS game), even small amounts of latency and jitter (as small as 60 ms) caused users to notice a drop in perceived quality.

The work examined in [5] attempted to group the games in prior studies based on model (avatar or omnipresent), perspective (first or third person or varied), and genre (FPS, Racing, Sports, RPT, RTS) to one of three levels of latency sensitivity. The thresholds of induced latency for high, medium and low sensitivity settings were estimated to be 100 ms, 500 ms, and 1,000 ms. More recent work [6] has addressed this

by exploring the effects of latency on classes of games. Using the open source third-person shooter game BZFlag, the authors illustrate that the range of tolerated latency can range by an order of magnitude (from less than 100 ms to 1.0 seconds) depending on the required precision for game play.

The work in [8] provides a useful framework for describing and assessing game quality of experience through the use of 'perceptual view inconsistency'. The framework assumes that 1) Quality of Experience (QoE) is based on an overall assessment that is based on individual (and decomposable) perceptions; 2) Inconsistencies result from the disconnect between the gamer's view and the server's (or global) view; 3) An overall QoE assessment requires each objective inconsistency to be mapped to corresponding subjective values. The authors refer to prior measurement work and show that the degradation of various metrics of perceived quality, each of which are specific to certain segments of specific games, can be modeled as a half normal (Gaussian) curve.

Based on the wide range of results that exists in the related work, we draw the following motivating conclusions:

- The impacts of latency and loss depend not just on the class of game, but also on particular modes of play and on specific gaming contexts.
- Establishing the relationship between latency and loss network measures with perceived quality is the most challenging issue that remains largely unsolved.

Our study develops a methodology for exploring the complex interaction between network impairment, the impacts on FPS online games, and the quality of experience perceived by end users (i.e., the gamers).

3 Methodology

3.1 Experimental Setup

Fig. 1 represents the experimental setup. For this setup, there are two sets of gameplayers. The first set consists of two players, referred to as "Participants", who play on a controlled network environment. Looking at Fig. 1, we emulate network impairment with the Linux *netem* capability (*Please refer to this site for further information on netem: http://www.linuxfoundation.org/collaborate/workgroups/networking/netem)* operating on our router. We add latency and jitter to packets arriving at the router destined for the 'congested participant'. The other participant is connected directly to the university's network and is referred to as the 'uncongested participant'. The second set of players, referred to as "Gamers", play from their home or dormitory room on campus during each session to maintain a controlled but realistic level of background competitors. The location of the game host is an experimental parameter. We designate either the congested participant or the uncongested participant as the host.

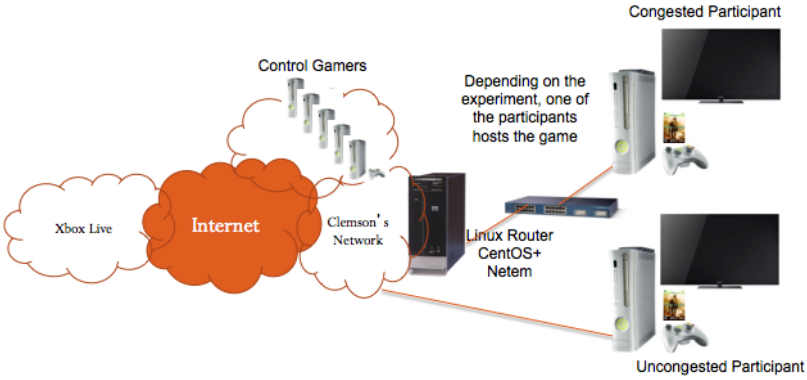


Fig. 1. Experiment Setup

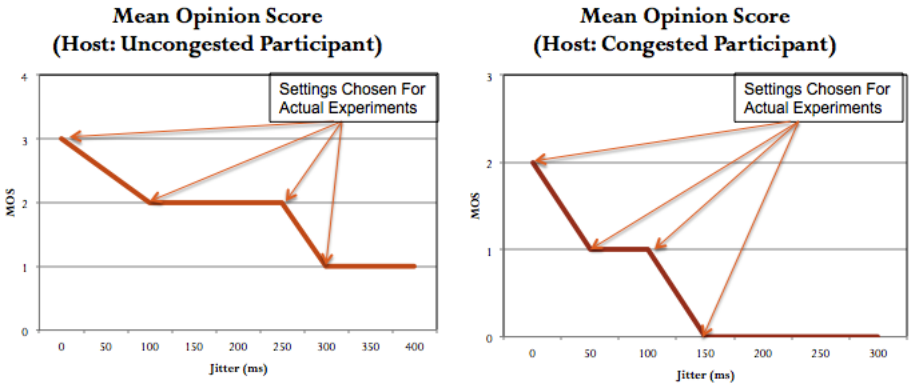
3.2 Experimental Procedure

Before each game session, the two participants are asked to fill out a pre-survey questionnaire geared to determine their expertise level and to estimate their network performance expectation based on their normal gameplay environment. Each game session lasts for ten minutes where all participants and gamers play in the "Free for All" game mode, and then the participants answer a post-survey questionnaire that consists of four Likert Scale survey questions. The four post-survey questions deal with each gamer's perception on following experiences: (i) Gameplay Satisfaction (ii) Gameplay Frustration (iii) Impact of Lag on Gameplay (iv) Likeliness to change network service providers. A MOS metric is computed for each participant using a weighted-sum approach where each of the four post-survey responses is given equal importance. The MOS metric derived for each participant is averaged to assess overall MOS (or QoE) for each network impairment setting. A detailed description of pre-survey questionnaire, post-survey questionnaire and MOS calculation is provided in [9].

3.3 Study Design

The study is divided into two phases: (i) Calibration phase (ii) Actual Experimentation phase. The 'calibration' phase involved a series of experiments designed to identify the set of network congestion settings that would be used in the actual study. The impact of jitter is quite different depending on which participant serves as the host. We conducted a pilot study in which we invited experienced Call of Duty Modern Warfare 2 players to play on several combinations of latency and jitter settings. In all experiments, 100 ms latency and a range of jitter in increments of 50 ms is added to all packets sent or received by the congested user. The results of the pilot study, presented in Fig 2, suggested that when the uncongested participant serves as the host, jitter in the range of [0, 300] ms is appropriate and when the congested participant serves as the host, jitter in the range of [0, 150] ms is appropriate. Moreover, the

functionality of *netem* was verified in the calibration phase of the study. A detailed description of these calibration phase findings is presented in [9]. Using the network congestion settings presented in Fig 2, a total of 56 experiments were conducted using university students as test subjects to obtain 7 MOS results per network congestion setting during the actual experimentations phase.



(a) (b)

Fig. 2. Pilot Study Results

4 Results and Analysis

We performed 56 total experiments, which resulted in 112 survey responses. Figs 3, 4, and 5 profile the participants. Fig 3 indicates that the participants are reflective of the undergraduate population at Clemson University. The majority rely on wired rather than wireless connectivity. Fig 4 suggests that the participants represent a broad range of FPS gaming experience, ranging from novice to expert. Finally, further analysis of the pre-survey responses suggests that 10% (about 10) of the participants were complete novices to any online games and another 10% were experts with Call of Duty Modern Warfare as shown in Fig 5. The rest of the participants had some level of experience with FPS games.

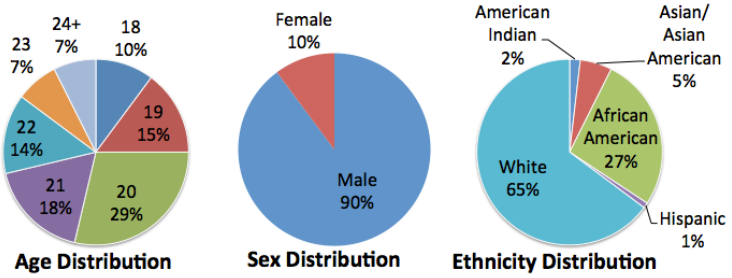


Fig. 3. Participant Demographics

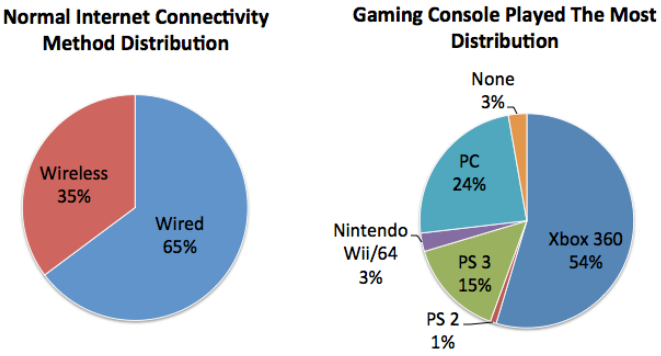


Fig. 4. Participant Internet Access Method

Participant Experience Level	Description
1	Plays 0-20 hours/week and Not FPS/MMOG games
2	Plays 0-20 hours/week of FPS/MMOG games or 21-40 hours/week of other games
3	Plays 21-40 hours/week of FPS/MMOG games or 41-60 hours/week of other games
4	Plays 41-60 hours/week of FPS/MMOG games or 61-80 hours/week of other games
5	Plays 61-80 hours/week of FPS/MMOG games or 81+ hours/week of other games

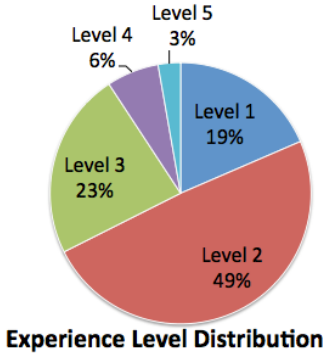
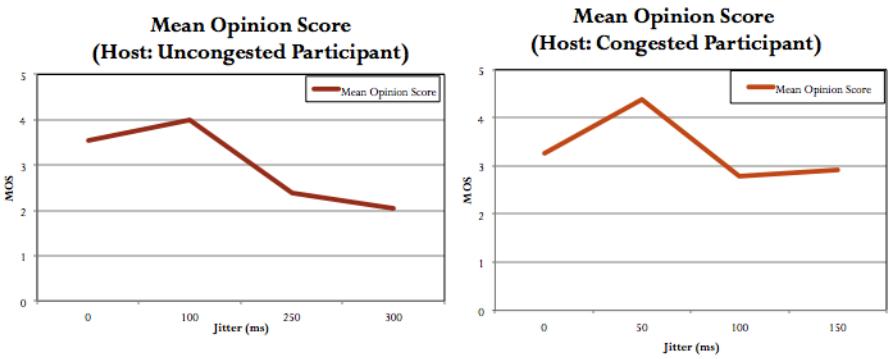


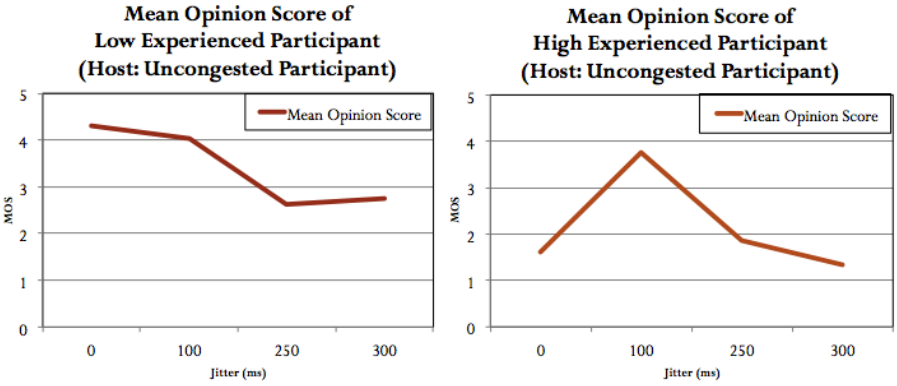
Fig. 5. Participant Gaming Experience

The results from actual experiments are shown in Fig 6(a) and 6(b) when the uncongested participant and congested participant are game hosts respectively. Note that only the results derived from post-survey responses of congested participants are presented in this study. The results derived from post-survey responses of uncongested participants are random in nature as these participants do not play under any network impairment and are presented in the extended version of this paper [9]. Fig 6(a) suggests that as jitter approaches 250 ms, the QoE degrades to a level (MOS < 3) that make the gamers' experience unacceptable when the game host is the uncongested participant. Fig 6(b) suggests that as jitter approaches 100 ms, the QoE degrades to a level (MOS < 3) that makes the gamers' experience unacceptable when the game host is the congested user. We observe that the MOS in Fig 6(b) drops as jitter is increased but not as definitively as in Fig 6(a). This happens because the congested participant is the game host in Fig 6(b). For this setting, the command and control data that the congested participant sends is actually avoiding the *netem* latency effects since the server is located at the local host. We conjecture that the outbound traffic stream is more sensitive to latency than the inbound traffic stream.

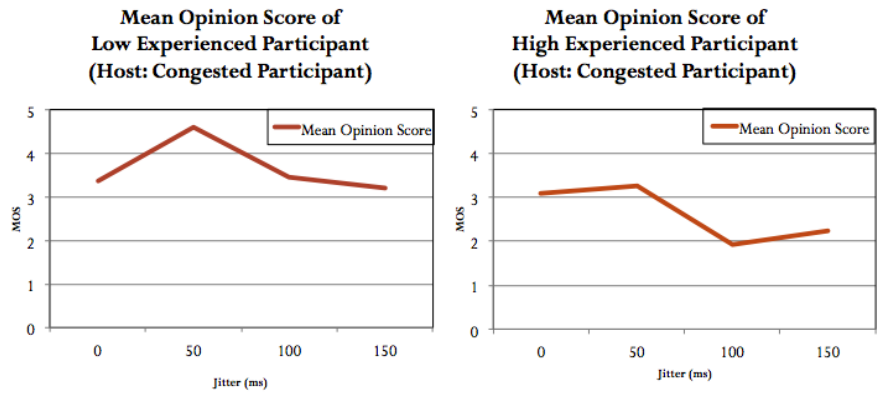


(a) (b)

Fig. 6. Overall Actual Experimentation Results



(a) (b)



(c) (d)

Fig. 7. Results based on Participant's Expertise Level

The results that compare the responses provided by experienced gamers vs. inexperienced gamers are presented in Figs 7(a)-(d). For both cases, when the game host is uncongested or congested participant, the results support our hypothesis by clearly suggesting that participant's level of expertise defines and determines the sensitivity to network congestion. The more experienced the participant (Figs 7(b) or 7(d)), the less tolerant he/she is to network congestion compared to the less experienced participants (Figs 7(a) or 7(c)). Also, the 'expert' or highly experienced gamers significantly berate even the lowest network congestion setting of 0 jitter/100 ms latency as seen in Figs 7(b) and 7(d) and have a hard time quantifying the level of congestion when the congestion settings are low.

5 Conclusions

We have studied the impact of latency and jitter on the perceived quality of the FPS game Call of Duty Modern Warfare 2. Our experimental methodology involved college students (gamers) playing a closed session of "Free for all" game mode. A total of seven gamers were involved in each game. Two gamers (referred to as participants) were under observation, one of whom had his/her network impaired in a controlled manner. The participants under observation provided a numeric assessment of their satisfaction, level of frustration, impact of lag and likeliness to change network service providers.

Our results suggest that, for the cases when the host is the uncongested user and the user under study is the congested user, a random jitter that is in a range of [0, 250] ms leads to a MOS less than 3. For the case when the host is the congested user and the user under study is the congested user, a random jitter in the range [0, 100] ms leads to a perceived MOS less than 3. Our original hypothesis was that *the perceived quality correlates to a gamer's skill level*. We did observe different thresholds of tolerance of network impairment based on the level of experience and skill. The more experienced the participant, the less tolerant he/she is to network congestion. The 'expert' or highly experienced gamers significantly berate even the lowest level of network impairment (with constant 100 ms latency and no jitter). However, these 'expert' gamers who are presumably the most sensitive to latency/jitter are also the most adept at compensating for impaired conditions as observed from their final game scores (which are not presented in this paper).

Our methodology has several limitations. First the number of user's in the study was not sufficient. While most data points in the results curves were based on 7 samples, some results (in particular those shown in Figs 7(a)-(d)) are based on only 2-3 samples. Second, our methodology did not precondition the test subjects. In other words, ideally, we would have a participant take our survey in his/her usual gaming environment. This provides a familiar context with which a user can base his/her assessment. In our next steps we would like to take a small set experienced Call of Duty gamers and first train their expectations and game play on 'ideal' conditions. Then, each gamer will periodically be asked their opinion allowing us to fully map study participants to their perceived quality based on controlled network settings.

References

1. Beigbeder, T., Coughlan, R., Lusher, C., Plunkett, J., Agu, E., Claypool, M.: The Effects of Loss and Latency on User Performance in Unreal Tournament 2003. In: Proceedings of 3rd ACM SIGCOMM Workshop on Network and System Support for Games, NetGames 2004 (August 2004)
2. Sheldon, N., Girard, E., Borg, S., Claypool, M., Agu, E.: The Effect of Latency on User Performance in Warcraft III. In: Proceedings of ACM Network and System Support for Games Workshop (NetGames), Redwood City, California, USA (May 2003)
3. Quax, P., Monsieurs, P., Lamotte, W., Vleeschauwer, D., Degrande, N.: Objective and Subjective Evaluation of the Influence of Small Amounts of Delay and Jitter on a Recent First Person Shooter Game. In: Proceedings of the SIGCOMM 2004 (August 2004)
4. Woundy, R., Lee, Y., Veiga, A., Williams, C.: Congestion Sensitivity of Real-time Residential Internet Applications. In: Proceedings of the SCTE Cable-Tec Technical Workshop (November 2010)
5. Claypool, M., Claypool, K.: Latency and Player Actions in Online Games. ACM Communications of the ACM (November 2006)
6. Claypool, M., Claypool, K.: Latency Can Kill: Precision and Deadline in Online Games. In: Proceedings of Multimedia Systems Conference (MMSys 2010) (February 2010)
7. Lang, T., Armitage, G.: A Ns2 Model for the Xbox System Link Game Halo. In: Proceedings of the Australian Telecommunications Networks and Applications Conference (2003)
8. Chan, P., Zarki, M.: Perceptual View Inconsistency: An Objective Evaluation Framework for Online Game Quality of Experience (QoE). In: ACM NetGames 2011 Workshop (October 2011)
9. Amin, R., Jackson, F., Gilbert, J., Martin, J., Shaw, T.: Assessing the Impact of Loss and Latency on the Perceived Quality of Call of Duty Modern Warfare 2, `http://people.cs.clemson.edu/~jmarty/projects/GamingAssessments/GamingStudyReportV6.0.pdf` (last accessed January 3, 2013)

The Effects of Negative Interaction Feedback in a Web Navigation Assistant

Marcelo G. Armentano and Analía A. Amandi

ISISTAN Research Institute (CONICET/UNICEN)
Campus Universitario, Paraje Arroyo Seco, Tandil (7000), Argentina
{marcelo.armentano,analia.amandi}@isistan.unicen.edu.ar

Abstract. Recommender systems are a common solution used to assist users in searching and retrieving information on the web due to the benefits that can be obtained from the evaluation and filtering of the vast amount of information available. This article presents a user study on the feasibility of using negative interaction, that is the absence of interaction with some items in a list of suggestions, as implicit feedback used to improve the performance of a web navigation assistant. Results showed an increment of 16.65% in the acceptance of the suggestions provided by the assistant and an increment of 43.05% in the average use of the suggestions window when using negative interaction with respect to not using this feedback mechanism.

Keywords: Intelligent Agents, Web Navigation Assistance, Implicit Feedback.

1 Introduction

The Web is daily queried with multiple purposes, such as reading the news, researching about a specific topic, planning vacations or searching for references to answer specific questions. Search engines facilitate finding information quickly on any topic of interest. Using a set of keywords as an input, search engines offer as a result a list of Web links hopefully addressing the issues described by those keywords.

Search engines offer a quick access to the information indexed from the web, automating the classification of the contents to deliver the most relevant information to each query. However, it is unusual that users are able find the information they intend to in their first search because they are constantly overloaded with content and links. This can be frustrating to users, who know that the information is somewhere available but they are not able to find it. A usual problem with search engines is that many users do not know how to "build" the search queries to obtain the information they are looking for. The queries used in a search engine are usually short and ambiguous, and different users might use the same query with completely different needs [1].

To address this problem we propose a web navigation assistant that recommends possible interesting web pages to the user while navigating the web. The assistant observes the user interaction with a web browser to obtain implicit indicators for the subject of the user's interests. Then, the assistant automatically performs new web

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 107–116, 2013.

searches to obtain pages related to those recently visited by the user. The interaction of the navigation assistant with the user is only through a small button indicating the availability of suggestions. When the user presses this button the top ranked results are shown in a window integrated with the web browser so that the user can continue navigating as usual if he/she is not interested in the suggestions.

We performed a set of experiments with real users in which we compared the results obtained considering the impact of two mechanisms of implicit feedback: the time spent on a page and the negative interaction. For these experiments we computed the percentage of usage of the window presented by the assistant, and the percentage of acceptance of the suggestions selected by the user. The results showed that considering both implicit mechanisms of feedback improve the performance of the assistant. Particularly, considering the negative interaction feedback allows the agent to modify his state when it has not properly inferred the user's search intention.

The rest of this paper is organized as follows. Section 2 presents some related work. Section 3 presents the web navigation assistant implemented, detailing its general architecture and how it process the implicit user feedback to improve the user profile for a searching session. Section 4 presents the user study conducted. Finally, Section 5 presents our conclusions of the study.

2 Related Work

There have been many research efforts focused on systems that recommend pages to web users. In this section we summarize some of those studies centered on obtaining implicit feedback from the user.

Our web navigation assistant learns a short term user model by analyzing the text of the pages that it assumes the user is interested in. Chi et al [2] obtains information about the user from the text of the links he/she clicks. However, many times the text of the links does not have relevant information about the linked page (e.g. "click here"). Zhu et al. [3] build a user model by analyzing the user's behavior in order to build rules such as "any word which is present in three consecutive pages visited by the user will be present in a page that is important to the user". These rules, called "abstract navigation patterns", are generated according characteristics of web pages visited by the user and not with any particular page. Matthijs and Radlinski [4] presented a personalization approach that combines the content and previously visited websites to build a user interest profile using the users' complete browsing behavior. This model is then used to re-rank web results. Hu et al. [5] also proposed a personalization approach but restricted to searching web services. They obtain the user interests from both the search requests and the previously used services by extracting the meaningful words from the service request and the used service description files.

The time spent in a page has been commonly used in many approaches as an indicator of the user interest in any piece of information. Parsons et al. [4], for example, found that there is a positive relationship between the time users spend watching an item available for purchase and their intention to buy it. Although the time spent watching an item depends also of external factors (such as the amount of visual

details, images, distractors, etc.), it is believed that it is a potential indicator of the user preferences.

Joachims et al. [7] studied the behavior of the users interacting with the results provided by a search engine. The results obtained shown that users take decisions from the descriptions provided by the search engine, by clicking on the links they are interested. However, the user clicks are conditioned by the order in which the links are shown. The user mainly interacted with the top ranked results (that is the links in the top of the list) even when the descriptions were less relevant that results in lower positions. Moreover, the user spent more time reading the descriptions of the top ranked results, and descriptions at lower positions received less attention.

Our approach takes the findings of Joachims et al. [7] to define the concept of negative interaction, that is to use the information of the search results skipped by the user to improve the user profile related to his/her information needs in a search session.

3 Web Navigation Assistance

WebHelper is a web navigation assistant we developed to observe the user behavior while interacting with a web browser to detect his/her subject of interest and suggest potentially interest pages. The objective of the assistant is to help the user to find web pages that he/she might not be able to find with the keywords he/she is using.

The search intention of the user is inferred by processing the text of the web pages he/she visit. After each visited page, the assistant updates the weights assigned to different words that describe the search intention of the user in a session. The top ranked words are then used as keywords to perform a new web search and present the results obtained to the user in an independent window integrated with the browser. This window is only shown when the assistant has enough information to infer user's the search intent. However, the user is able to close this window at any time and to open it asking for suggestions.

3.1 General Architecture

WebHelper is implemented server-side, in the form of a web proxy. This way, the user can start using the assistant simply configuring his/her web browser to use the assistant's proxy. When the user sets his/her browser to use this proxy, all requests are intercepted and two processes are applied. The first process extracts the information contained in the response pages and assigns different weights according to the context of each extracted term. The second process injects in the HTML response the code corresponding to the suggestion window, if appropriate.

The dataflow in *WebHelper* is shown in Fig 1. The user searches the internet using a web browser. The *WebProcessor* module captures the response to the server request and executes two processes: information extraction and information injection. The information extraction process captures the content of the web page and passes it to the *WordProcessor* module. If the assistant has any recommendation to present to the

user, the information injection process adds to the response the code corresponding to the suggestions window to be presented to the user. On the other hand, the *WordProcessor* module takes the page content received from the *WebProcessor* module and updates the user profile. We describe this process in Section 3.2

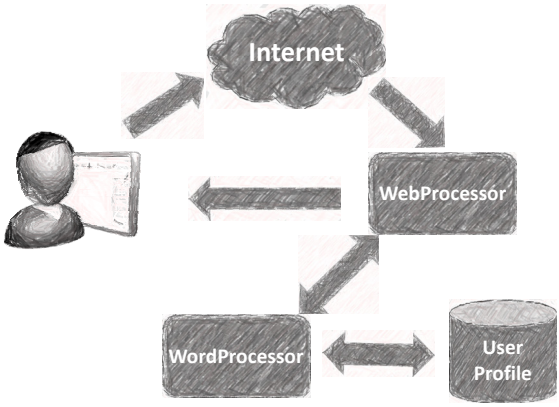


Fig. 1. WebHelper general architecture

3.2 Information Processing

When the *WordProcessor* component receives a new visited document, it first converts it to plain text, detects its language[1], removes stop words and applies the corresponding Porter stemming algorithm [10].

The user profile keeps the set of most frequent words, each of them with an associated weight. These weights vary according to the place in the page in which each word appears, and vary after each visited page, either by adding or subtracting a new constant value. Different constant values were defined for words appearing in headings, normal text, links, text typed in input boxes, meta-information, etc.

On the other hand, words weights in the user profile are decremented for three different reasons: (1) the word does not appear in the current page; (2) the negative interaction feedback is applied; (3) the user does not select any suggestion after a fixed amount of visited pages. We give details about the last two items in Section 3.3.

After processing each visited page, the following situations can arise:

- There were no suggestions up to that moment, and no terms in the user profile exceed the "suggest-me" threshold. In this case, nothing happens and no suggestions are given to the user.
- There were no suggestions up to that moment, and two or more words exceed the "suggest-me" threshold. In this case, the assistant use the words exceeding the threshold and performs a web search using those words as keywords. The top results of the search engine are presented as suggestions. In the current implementation of *WebHelper*, the Google® API is used for querying the web.

[1] We currently work with English and Spanish

- The assistant was already showing suggestions, but after processing the information of the new document the weights of all words are under the "suggest-me" threshold. In this case, the suggestions window is not injected in the response to the user.
- The assistant was already showing suggestions and after processing the information of the new document the words whose weights exceed the "suggest-me" threshold are the same that in the previous cycle. In this case, suggestions are rotated, and a new set of web pages from the previous query are shown in the suggestions window.
- The assistant was already showing suggestions and after processing the information of the new document a new set of words weights exceed the "suggest-me" threshold. In this case a new web search is performed using those new words as keywords and the top results are presented as suggestions, replacing the previous window.

Before querying the web, the Google® API is used to obtain suggestions for each candidate keyword. This service provided by the Google® API, receives a word as input and returns a set of related words as output. The suggested words for each keyword and the keywords themselves are combined to query the search engine. This step improved the performance of the assistant since the terms of in the user profile are words trimmed by the stemming process. Using those stems to query the search engine produced unexpected results.

3.3 Feedback Processing

To determine the search intention of the user in a session, the navigation assistant considers two implicit feedback indicators: the time the user spends in a page and the interaction with the suggestions window.

The time the user spends on a page have been extensively used as an indicator of interest in a piece of information [4][8]. A web page is considered active by *WebHelper* when the user first access to it or when it is selected by using the browser's tabs.

We consider two types of interactions with the suggestions window: positive and negative interactions. A positive interaction is triggered when the user clicks on link in the suggestions window. In this case, the content of the suggestion (title and snippet) is used to improve the user profile for that session.

Moreover, at the same moment when the user shows interest in a suggestion, there is also a negative interaction with respect to the items located above the one selected by the user in the list of suggestions. We apply a heuristic based on statistical studies that showed that lists are usually read from top to bottom [7]. This way when the user selects an item in a list, he/she is not only indicating interest in the suggestion selected but also is implicitly indicating that he/she is not interested on the previous items. Assume for example that the assistant presented a list of suggestions <s1, s2, s3, s4, s5, s6, s7> and the user clicked on s1, s3 and s5. We can assume that s3 is more relevant that s2, since the user read the description of s2, but preferred clicking on s3.

For the same reason we can assume that s5 is more relevant that s2 and s4. We can assume that s2 and s4 are not relevant to the user's needs and we can use the information they contain to update the user profile as negative information [9]. The suggestions list can then be re-organized to eliminate the preceding recommendations.

Similarly, if the user has not interacted with the suggestions window after a certain amount of visited pages, it is assumed that the suggestions are not interesting to the user. Consequently, suggestions are rotated to show the user a new set of possible interesting websites. Following the idea of negative interaction, when suggestions are rotated the weights of the terms appearing in the hidden suggestions information are decreased.

3.4 Suggestions Window

When the assistant has suggestions to make to the user, it creates a drop down window integrated to the page returned to the user. This window is situated on the right of the returned page and contains a maximum of eight suggestions, consisting in a title and a snippet (Fig. 2). This window can be hidden and re-opened at any time.

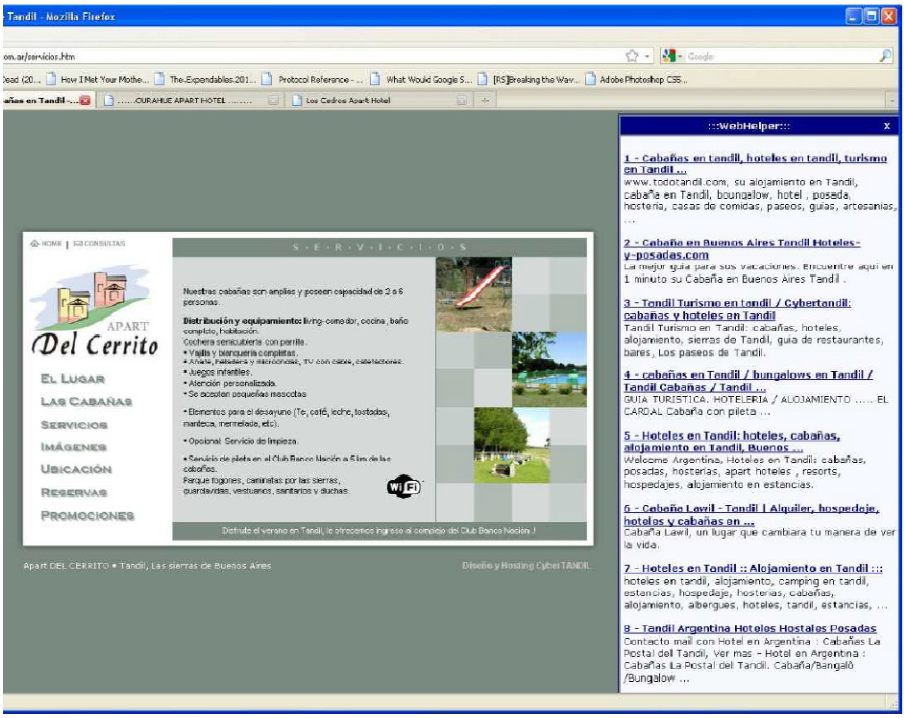


Fig. 2. Suggestions window while visiting a page

4 Experimental Evaluation

4.1 Experiment Setup

To validate the proposed approach, we tested our system with 32 volunteer users, both male and female, with ages ranging 18 to 55, and different levels of expertise in searching the web. Users were divided into groups to test four different configurations of the system.

We asked each user to freely use the web to search for information about any subject of interest. The navigation assistant was slightly modified for these experiments: after selecting a suggested site from the suggestions' window, users were asked whether the page they had just visited was interesting according to their search intentions. This feedback was only logged to compute the performance of the assistant and it was not used to modify its normal behavior since in a live scenario feedback is never asked to users.

The duration of the experiments was, on average, 25 minutes. Most users started their session accessing Google's search engine and entered the keywords they believed best expressed their information needs. Other users interested in news opened their preferred online newspaper and start browsing through it. Finally, other users visited pages they usually visit.

4.2 Metrics

Two metrics were used to evaluate the performance of our navigation assistance: the percentage of use of the suggestions' window and the percentage of acceptance of the suggestions selected by the user.

To compute the average use of the suggestions' window, we divided the number of suggestion selected by the user by the number of windows shown to the user during the experiment.

$$use\_of\_the\_assistant = \frac{\#suggestions\_selected}{\#suggestions\_shown}$$

Likewise, the percentage of acceptance of the suggestions is computed dividing the average number of accepted suggestions by the average number of suggestions selected by the user.

$$acceptance = \frac{\#suggestions\_accepted}{\#suggestions\_selected}$$

4.3 Time Spent on a Page

The first experiment explored whether considering the time the user spent on a page any influence the performance of the assistant. In the first configuration of the navigation assistant, any page visited by the user was processed, without considering how

long the user remained in the same page. We use this first configuration as a base-line for comparison with the other three configurations. In the second configuration, pages were only considered for improving the user profile when the user spent more than a predefined amount of time reading the page. For our experiments this threshold was set to six seconds. In this configuration, the system took more time to generate suggestions since not all information visited by the user is processed and, consequently, the weights assigned to individual terms increased more slowly than in the first configuration.

Since different experiments took different durations, the average results were normalized by the duration of the individual experiment. For example, nine suggestions windows were shown to User 1 for the first configuration, and User 1 used the system for 30 minutes, so we used a value of 0.3 windows per minute. Table 1 shows the results obtained.

Table 1. Results for user tests with and without time processing

	Suggestion windows shown	*Suggestions selected by the user*	*Suggestions accepted by the user*	*Time for showing the first suggestions*	*Duration of the experiment*	*Google searches*
Without Time processing	0.662 per minute	0.148 per minute	0.113 per minute	3.62 minutes	22.125 minutes	0.414 per minute
With time processing	0.350 per minute	0.184 per minute	0.163 per minute	10.62 minutes	22.375 minutes	0.373 per minute

From Table 1 we can compute that the average use of the assistant without considering the time spent on a page was 22.35% while the average acceptance of the suggestions selected by the user was 76.35%. On the other hand, when considering the time the user spent on a page, these values are increased to 52.57% and 88.59% respectively.

4.4 Negative Interaction Feedback

The third configuration of the experiment was set up to consider negative interaction feedback but did not consider the time spent on a page. Finally, the fourth configuration used both feedback mechanisms together. Table 2 shows the results obtained. A new column is added to show the number of times that the suggestions were rotated.

The average acceptance of the suggestions selected by the user was 87.1% (with an average use of 26.59%) if we consider the negative interaction feedback alone. However, we reach an average acceptance of 93% and an average use of 65.4% by combining both feedback mechanisms.

Table 2. Results for user tests with and without time processing

	Suggestion windows shown	*Suggestions selected by the user*	*Suggestions accepted by the user*	*Time for showing the first suggestions*	*Duration of the experiment*	*Google searches*	*Rotation of suggestions*
Without Time processing	0.583 per minute	0.155 per minute	0.135 per minute	8.750 minutes	24.000 minutes	0.307 per minute	0.102 per minute
With time processing	0.327 per minute	0.214 per minute	0.199 per minute	10.500 minutes	22.625 minutes	0.140 per minute	0.072 per minute

4.5 Discussion

For the first configuration, while the acceptance of the suggestions selected by the user was high (76.35%), the use of the suggestions window was quite low (22.35%). This fact suggests that the information shown to the user in the suggestions' window might have not convinced the user to click on a suggestion. Since for this experiment all visited pages were processed, the threshold for showing suggestions was reached quickly, but the system was not able to properly infer the intention search of the user. The second experiment solved this problem by increasing the use of the suggestions window to 56.57% at the expense of the time that the system needed to give assistance to the user.

On the other hand, by using the negative interaction feedback alone, we obtained an acceptance of the suggestions similar to that we obtained when we used the time processing alone. The average usage in this case was better than the first configuration but lower than the second configuration. The time spent to show suggestions to the user is lower than the second configuration, but higher than the first configuration. This is due to the fact that the negative interaction lowers the weights of the terms that are present in the unvisited documents. Then these weights trend to increase slowly.

By combining both feedback mechanisms we obtained a better general performance of the assistant. Furthermore, although the time that took the system to show the first suggestions window was similar to that in the second configuration, it performed fewer queries to the search engine, achieving a better acceptance of the suggested sites.

5 Conclusions

In this article we described a web navigation assistant that infers the search intention of a user in order to provide him with suggestions of sites in which he/she might be interested. The assistant uses implicit feedback mechanisms to improve a temporal user profile that contains information about the subject of interest in a search session.

We presented a user study on the feasibility of using negative interaction, which is the absence of interaction with some items in a list of suggestions, as implicit feedback used to improve the performance of the web navigation assistant.

The study allowed us to conclude that using both the time spent on a page and the negative interaction feedback improves the performance of the assistant, obtaining an increment of 16.65% in the acceptance of the suggestions provided and an increment of 43.05% in the average use of the suggestions window.

References

1. Hu, R., Dou, W., Liu, X.F., Liu, J.: Personalized Searching for Web Service Using User Interests. In: Proceedings of IEEE Ninth International Conference on the Dependable, Autonomic and Secure Computing (DASC), pp. 156–163 (2011)
2. Chi, E., Pirolli, P., Chen, K., Pitkow, J.: Using information scent to model user information needs and actions and the Web. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 490–497. ACM, New York (2001)
3. Zhu, T., Greiner, R., Häubl, G.: Learning a model of a web user's interests. In: Brusilovsky, P., Corbett, A.T., de Rosis, F. (eds.) UM 2003. LNCS, vol. 2702, pp. 65–75. Springer, Heidelberg (2003)
4. Matthijs, N., Radlinski, F.: Personalizing web search using long term browsing history. In: Proceedings of the Fourth ACM International Conference on Web Search and Data Mining (WSDM 2011), pp. 25–34. ACM, New York (2011)
5. Hu, R., Wanchun, D., Liu, X., Liu, J.: Personalized Searching for Web Service Using User Interests. In: Proceedings of the 2011 IEEE Ninth International Conference on Dependable, Autonomic and Secure Computing (DASC 2011), pp. 156–163. IEEE Computer Society, Washington, DC (2011)
6. Parsons, J., Ralph, P., Gallagher, K.: Using Viewing Time to Infer User Preference in Recommender Systems. In: Proceedings of the AAAI Workshop on Semantic Web Personalization held in conjunction with the 9th National Conference on Artificial Intelligence (2004)
7. Joachims, T., Granka, L., Pan, B., Hembrooke, H., Gay, G.: Accurately interpreting clickthrough data as implicit feedback. In: Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2005), pp. 154–161 (2005)
8. Konstan, J., Miller, B., Maltz, D., Herlocker, J., Gordon, L., Riedl, J.: GroupLens: applying collaborative filtering to Usenet news. Commununications of ACM 40(3), 77–87 (1997)
9. Kurosu, M., Ookawa, Y.: Effects of Negative Information on Acquiring Procedural Knowledge. In: Proceedings of the International Conference on Computers in Education (ICCE 2002), p. 1371 (2002)
10. Porter, M.: An Algorithm for Suffix Stripping. Program 14(3), 130–137 (1980)

Automatic Layout Generation for Digital Photo Albums: A User Study

Francine B. Bergmann, Isabel H. Manssour, Milene S. Silveira, and João Batista S. de Oliveira

Faculdade de Informática, Pontifícia Universidade Católica do Rio Grande do Sul
Avenida Ipiranga 6681, 90619-900, Porto Alegre, RS, Brazil
francine.bergmann@acad.pucrs.br,
{isabel.manssour,milene.silveira,joao.souza}@pucrs.br

Abstract. The low cost and ubiquitousness of digital cameras allow ever larger collections of photos. Although storage is not usually a problem, selection for presentation requires time and effort. Studies show the existence of user groups that never tried to create digital photo albums or were discouraged by the difficulty of this task. Trying to minimize the effort to create albums, we developed a prototype called Twingle. Using Twingle the creation of an album requires dragging pictures into a page and the program is responsible for positioning and resizing. To investigate whether such an automatic layout generation could be an alternative to users, we made a study exploring Twingle with users from several ages and professional areas. The related works, the performed user study and its results are discussed here.

Keywords: digital photo albums, automatic layout generation.

1 Introduction

Before the age of digital cameras, several considerations limited the amount of photos taken by most people and affected the practices adopted for their organization and presentation: a small number of photos were taken and developed to be organized as photo albums. The access to digital cameras favored the growth of photo collections [2,8,11], and these much larger collections demand new tools and practices for their organization.

It was observed [4] that even though the most accepted way of viewing pictures are the printed albums, most people do not create them. Some never tried, and some of those who tried were discouraged by the difficulty of creating an album [4]. Users of commercial tools consider the creation of digital photo albums (DPA) a costly and time consuming task [20], since these tools require manual edition of the pages by placing pictures, text, backgrounds and decorative elements [17].

In this context, we believe that the automatic layout generation of DPAs (or part of them) could benefit users who have difficulties to perform this task or do not want to spend time at it. To test this, it is necessary to explore and understand the behavior of

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 117–126, 2013.

users of such tools and find out the needs that these tools should address. In this context, we developed a prototype called Twingle that automates the process of DPA production using automatic layout generation. Twingle tries to minimize the effort to create albums, since it only requires dragging pictures into a page and the program is responsible for positioning and resizing them.

In order to investigate whether such an automatic layout generation could be an alternative to users, we made a study with people from several professional areas and ages exploring Twingle. The results and a discussion about our findings, as well as the related works, are presented in the following sections.

2 Related Works

Most people are photographed or receive photos at some point in life [18]. In this context, several authors [4,6,7,14,16] tried to understand how people deal with their photo collections and those studies show an evolution in the practices adopted during the transition from non-digital to digital photos.

Photos organized in an album usually tell a story and people feel obliged to organize their pictures for family and friends. When they do not do that, they say that "ought to" or "definitely planned to" [18]. Families usually select and arrange their favorite photos into albums to present them personally to friends and relatives, but the production of photo albums is considered complex and time consuming [4]. The same study also stated that few people spent time organizing their digital photos [4], but they usually "throw" their photos into a folder similarly to a shoe box where pictures are often stored [14].

Another study shows that most people organized their digital photos into directories according to the convention "date/event" or similar [7]. To do this, they use interfaces similar to Windows Explorer, rejecting photo management software. A similar study [1] shows that people place images into directories and feel more satisfied with the organization of their digital photos than the non-digital, as they had "all in one place", giving an illusion of organization, since the digital photos do not take up physical space.

To ease the organization and creation of photo albums, several studies [1,8, 11,13,14,16] have proposed methods and tools to manage large collections of images, integrating resources such as timestamps, GPS data and face recognition information as a way of clustering photos into collections. The goal is to speed up and ease the user experience in time-consuming and tedious tasks [1].

In addition, to understand how people deal with their photo collections, several works [1,5,6,9,16,21,22] focused on the creation of DPAs, and tools [10,19] have been developed commercially to the same purpose. DPAs (or photobooks) are typically produced for printing. The result is usually a file to be sent to the owner of the software, which produces a high-quality printing of the album. These tools (available online or in desktop versions) allow for the free edition of the layout and use templates [10,19] to organize the images on the page. Others [3,9,20] provide solutions for the generation of collages, solve problems of unfilled spaces in album pages, and

generate DPAs with only a few initial selections required. There are still other authoring systems that automate user tasks [5,22] like image grouping, background selection, automatic page layout and other functions.

3 User Study

To investigate whether a simple interactive tool would ease the production of photo albums, helping to assuage the lack of time and knowledge of the users, a prototype called Twingle - characterized by the automatic organization of the pictures and texts in photo albums - was offered to a group of people.

The study had two phases: the first explores the first version of Twingle. In this phase, several enhancement suggestions were given by users to improve their experience with the production of DPAs. The second phase was performed with an improved version of Twingle, containing the most relevant functionalities suggested. Subsequent sections describe Twingle and the two phases of the study.

3.1 What Is Twingle?

Twingle is a prototype that tries to reduce the effort to create photo albums: the user does not need to position or resize images on the page and the layout is done automatically. It was implemented using HTML/CSS and JavaScript, allowing its use in any web browser supporting these languages, and communicates with a server responsible for the layout algorithm.

The creation of an album requires only dragging images into a page, as the program on the server is responsible for positioning and resizing images using an algorithm for automatic layout generation [12]. Briefly, the algorithm receives a set of rectangular items to be placed on the page and estimates the amount of area needed by each item, using a divide and conquer approach to put them all on the page. To each added item, the procedure is performed again to position it in a visually pleasant way, in which images and their assigned regions should have similar aspect ratios. Moreover, the page needs to be homogeneously covered, that is, items should be evenly placed on the page and not cluttered on some region, not leaving excessive empty space and filling the entire page as much as possible.

There are several tools and systems mentioned on Section 2 that may be similar to Twingle, however none of them used an automatic layout generation for DPAs composed of texts and images selected by the user. Works that provide an algorithm for automatic layout generation [3,9,20] do not support text inclusion, whereas some commercial tools allow the creation of DPAs, but only using templates [10,19]. Others [1,5,22] have different purposes, automating the task of selecting photos or allowing a different way of interaction [6] or organizing the pictures [16]. Thus, Twingle was designed with a specific goal to investigate the reasons behind the difficulty of producing and presenting DPAs.

3.2 First Phase

In the first phase, besides dragging images to compose a page, Twingle offered the following features: start a new album, add new pages, view the album in full size, generate a PDF file for printing, as well as delete an image from the page, add a caption to a picture or shuffle the images to obtain a new page layout. Figure 1 shows a sample album.

Fig. 1. Page generated by Twingle as images and captions are inserted

Objective.
To study whether a simpler tool for page design would be well received by users. A prototype was tested to produce DPAs just by dragging and dropping images to a page, and layout was produced automatically.

Description.
This study was conducted with 15 users from several professional areas, such as computer science, graphic design, architecture, management and psychology, among others, and ages between 18 and 60 years.

A pre-interview showed that not all users like to make photo albums, and some of those who like sometimes thought the task was too laborious and time consuming. Thus, these users were organized into three groups (five users each), according to their willingness to create DPAs: Group 1, users that like to make photo albums; Group 2, users that would make photo albums if they knew how or if it were "practical"; and, Group 3, users that do not like/do not wish to make photo albums.

To split the users into such groups, their answers to the following four questions were analyzed: Generally, how do you present photos to your friends/family? Do you like to make photo albums? Do you make traditional photo albums (print and place photos in an album)? What is your opinion about making DPAs?

Each user had the opportunity to explore the software and produce albums with it. As the purpose of the study was not to evaluate the tool but to allow users to evaluate a simpler process of automatic creation of DPAs, they were free to ask questions to the observer during the study.

Thereafter, an interview was made with open questions (the same for all groups) and one variable question according to the specific group. The main idea was to identify, during the experience with Twingle, the users' behavior and needs regarding automatic tools for creating DPAs.

Discussion.

Each user was identified by a group number (1 to 3) and a letter (A to E). Among the 15 users, 9 showed previous discouragement about the creation of photo albums due to time and work involved, including some of the users who like to make photo albums, as the case of user 1D, which considers commercial tools too laborious for people not used to designing documents: "*The matter of design is troublesome. As I am not a designer, it took me a month doing an album with another tool, changing the orientation and photos on it.*".

Considering these users - who like to make albums - three of them are used to tools for DPA generation. When asked if they would make an album using Twingle, they said they possibly would, but more features need to be available. On the other hand, those who had never used a tool to create albums said that they could use Twingle, as mentioned by user 1C: "*I really liked, it is very practical. It's easier than the programs I normally use to print photos, such as Word. I tried to organize everything into Word or printed directly from Paint, Photoshop.*".

Among those in group 2, 2C uses a tool for DPA generation, and 2B did not know these tools; the others knew but did not use them. After the experience of using Twingle, only 2A had restrictions about its use, saying that he would use it only if additional resources were available.

In group 3, only two users previously used a tool for DPA generation: 3A and 3D indicated that it is a laborious task and 3A also highlighted that when using the automatic layout, the results were unsatisfactory: "*I take a long time, until today I used only one automatic tool, but the result was disastrous.*".

After using Twingle, when asked about its use for generating albums, 3D said he would not use it because he prefers to have full control over the final layout: "*I am a user who likes to be in control (...). I like to adjust the size of the photos or place of the caption (...) I like to move items to let it the way I want (...).*". User 3A says he would make those without restrictions, and the others would require more features.

When asked whether the results of Twingle could be shown to friends and relatives, only two users had restrictions to the final layout: 1A and 1B, which usually make albums. User 1A indicates: "*As it is today I don't think so, it lacks some features.*", and user 1B said that he preferred to show the photos in full screen. Among the main reasons for presenting the results of Twingle for friends and family, the practical organization of the photos and the pleasing results were mentioned by users from the three groups.

When asked whether they would print the results, there was a trend to use the result in the computer instead of printing albums, and users mentioned emailing or publishing on the internet. Some of the reasons for not printing are the absence of resources they would like and the lack of knowledge about places for printing a PDF file with photo quality.

User comments show that for better results some features are especially important. Among the most mentioned functionalities, were:

- change the position of a picture while keeping the general layout;
- resizing a specific photo;
- choose the page orientation (portrait or landscape);
- have background or ornament effects available;
- be able to place general texts (not captions) on the page;
- free control of the spacing between images;
- preview the results on the screen, including video and interaction.

3.3 Second Phase

To the second phase, Twingle was refined according to the results from the first phase. The most relevant features demanded by users were added: upload (working both for photos and backgrounds), addition of general text/headlines, page orientation, background insertion and interactive possibilities to resize or change the position of a photo, and delete a whole page. Figure 2 shows a DPA created with the new version, containing a headline, several photos, text areas and captions.

Fig. 2. DPA album generated by new version of Twingle

Objective.
This phase was performed to verify if there was any difference in the results compared to the ones from the first phase and to check if the implementation of the new features had turned the experience of creating photo albums more enjoyable and less complex.

Description.
This phase was performed 8 months after the first one. Participants belong to the same professional and age groups mentioned previously.

To determine whether there was an improvement of the user experience, we called new participants, as well as volunteers from each group from the previous research. The methodology used and the classification of the new volunteers into groups

remained the same from the previous phase. We had the participation of 8 people from the first phase, and 6 new volunteers, totaling 14 participants: five users at Group 1, five at Group 2 and four at Group 3.

The new participants received the same questionnaire of the previous phase according to their classification into the groups, and the old ones answered a new list of questions, among which we mention: What did you think of the new features of Twingle? On what occasions would you use this tool?

Discussion.

For analysis and discussion of the results, the 8 volunteers from the first phase were identified by their previous codes (A to E) and the 6 new users were numbered in sequence (F to G).

Regarding to the new functionalities, 6 from the 8 participants of the first phase acknowledged that both the use experience and the results provided by the new version were better, and 11 from the 14 users said it was "*quite good*".

Among the 6 new users, 3 never used photo album software and believe such work is tiresome and time-consuming. When Twingle was used for the first time they observed that the process may be quick and practical, and showed interest in using such a tool to organize their pictures for family and friends. Participant 2F said: "*I did not know those tools, I had no idea how easy it was to produce an album. It is more interesting to see the pictures like this, as I use to organize them by event and it would be nice to put them all on the same page.*". 2G talks about Twingle as an image organizer: "*It's a very simple way I have to organize my photos. I found it very practical, fast, and creating the album was an easy task.*".

Four participants suggested a new feature for Twingle: to generate pages automatically without the user having to drag the images individually, but just select a folder on their computer and let the program take care of producing the album. User 3A says: "*Since this is an automatic tool, it is interesting that I could insert all my photos and the program returns the pages of the album.*".

Among the 4 users from the third group, those that dislike or do not consider producing albums, only one continued disliking after using the new version of Twingle. The reason was that this user does not enjoy taking pictures or manipulating them digitally, and does not even publish them in social networks. That user said that he would print the pages and send them to friends and relatives, but only if it was really necessary and if the results were pleasant.

On the other hand, people used to full-featured album tools missed more capabilities. User 1B said: "*It is far from what I want from album software, as this layout is not enough. I wish to place pictures in different positions and change their sizes freely.*".

Regarding the new features, the possibility to use backgrounds was valued by users, and 7 from the 14 users wished to add titles to the top of the page, to be handled differently from the current model (where texts are placed in a box and positioned as images) and to be manipulated outside the reach of the automatic layout algorithm.

Suggestions for changes among the participants concerned page margins, text fonts, and captions, as well as their color and alignment. Swapping the position of two images was also suggested. When asked about the lack of features of the tool, only 2

from the 14 users found that Twingle needed more functionalities. The participant 3G says: "*I would use [the tool] to print photos, and for doing that I think the features it has are enough.*".

3.4 Final Analysis

The study found a tendency of people to increase the use of digital content and, therefore, a need for ways to publish photo albums in digital format, both as PDF and HTML files.

Some users were receptive to the idea of a tool for making albums in a fast and simple way. In the first phase, 11 from the 15 users who do not use similar tools began to consider the possibility of using them, and respectively 10 from the 14 users of the second phase, especially because of the increased speed given by the automatic layout.

In the second phase 70% of the participants, who are frequent users of social networks, were motivated to use Twingle if there was an application for Facebook, revealing an interesting proposal to be investigated.

With respect to functionality, the amount of suggestions for improvement on the second phase fell considerably compared to the first version. This leads us to believe that the options present in the second version of Twingle could be in accordance with the requirements of some users, making the user experience more usable and satisfactory, and providing a more pleasing result.

Even with the evolution of Twingle, people who like to create photo albums continued feeling the lack of editing features, listing several new options to change the tool and making Twingle deviate from its initial purpose. In this case, perhaps Twingle's layout algorithm may become an add-on to programs already on the market, as an alternative capable of providing a solution for those who do not have time or patience to spend on making photo albums.

The items described in this final analysis highlight the potential of automated tools such as Twingle. While some people who enjoy all possible resources present in commercial tools, there are also those who need or prefer to design a DPA quickly and easily. Thus, it is essential to have an option to facilitate the construction of DPA while providing a pleasant result that does not require further edition. These tools should be capable of generating pages without the user having to manually select the photos, but only indicating a folder containing images he considers interesting to the album to be produced. The main intention is to reduce labor and time required to create DPAs. Therefore, through Twingle, we notice that there is room for tools for different user profiles, from those who enjoy more powerful features for editing, to those who just want to point to a folder on the computer and have their album generated.

4 Conclusions and Future Work

In this work a literature review showed that it is not uncommon for people to be discouraged to create photo albums because they consider it a costly and time consuming

task. Thus, we investigated, through a user study, if the automatic generation of DPA would reduce the effort.

The results of the study show the potential of the automatic layout generation in DPA applications. While some people enjoy controlling all possible resources of commercial tools to edit the layout of DPAs, others prefer a faster and easier alternative, producing DPAs through less options and reduced effort. Thus, it is essential to have an option to facilitate the construction of DPA while providing a pleasant result that does not require further (or too much) edition. The main aim is to reduce issues such as excessive labor and time pointed out by the people who need or are used to create photo albums.

All users were enthusiastic about the possibility of generating automatic layout, but they highlight features that a tool like this should provide: change the position of a picture while keeping the general layout; photo resizing; choose the page orientation; provide background ornament effects; allow the placement of general texts on the page; provide free control of the spacing between pictures and the possibility of inserting several photos at once.

After the user study, it was possible to verify that users were receptive to the idea of a tool for making albums in a fast and simple way. More than 71% of them who do not use similar tools began to consider the possibility of using them, especially because of the increased speed and good results given by the automatic layout.

Currently we are developing a version of Twingle for Facebook, and exploring the potential of its automatic layout generation for other applications, such as picture panels and portfolios. In addition, we intend to investigate the use of tools like Twingle in tablets, to identify alternative features made possible by the different forms of interaction available on these devices.

Acknowledgements. The results presented in this paper were obtained in cooperation with Hewlett-Packard Brazil Ltda., using the Brazilian Informatics Law (Law nº 8.248 of 1991).

References

1. Boll, S., Sandhaus, P., Scherp, A., Westermann, U.: Semantics, content, and structure of many for the creation of personal photo albums. In: Proceedings of the 15th international conference on Multimedia (MULTIMEDIA 2007), pp. 641–650. ACM, New York (2007)
2. Cunningham, S., Masoodian, M.: Identifying personal photo digital library features. In: Proceedings of the 7th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL 2007), pp. 400–401. ACM, New York (2007)
3. Diakopoulos, N., Essa, I.: Mediating photo collage authoring. In: Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology (UIST 2005), pp. 183–186. ACM, New York (2005)
4. Frohlich, D., Kuchinsky, A., Pering, C., Don, A., Ariss, S.: Requirements for photoware. In: Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work (CSCW 2002), pp. 166–175. ACM, New York (2002)

5. Gao, Y., Atkins, C., Cheatle, P., Xiao, J., Zhang, X., Chao, H., Wu, P., Tretter, D., Slatter, D., Carter, A., Penny, P., Willis, C.: MagicPhotobook: designer inspired, user perfected photo albums. In: Proceedings of the 17th ACM International Conference on Multimedia (MM 2009), pp. 979–980. ACM, New York (2009)
6. Jin, Y.K., Choi, S., Chung, A., Lee, I.M., Kim, M.C., Woo, J.: GIA: design of a gesture-based interaction photo album. Personal Ubiquitous Comput. 8(3-4), 227–233 (2004)
7. Kirk, D., Sellen, A., Rother, C., Wood, K.: Understanding photowork. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2006), pp. 761–770. ACM, New York (2006)
8. Lee, B.N., Chen, W., Chang, E.Y.: Fotofiti: web service for photo management. In: Proceedings of the 14th Annual ACM International Conference on Multimedia (MULTIMEDIA 2006), pp. 485–486. ACM, New York (2006)
9. Li, X., Takayama, Y., Shimizu, N., Mori, T.: An Aesthetic Solution for Photo Collection Page Layout. In: International Conference on Knowledge and Systems Engineering (KSE 2009), pp. 33–38 (2009)
10. MyPublisher, http://mypublisher.com
11. O'Hare, N., Gurrin, C., Lee, H., Murphy, N., Smeaton, A.F., Jones, G.J.F.: My digital photos: where and when? In: Proceedings of the 13th Annual ACM International Conference on Multimedia (MULTIMEDIA 2005), pp. 261–262. ACM, New York (2005)
12. Oliveira, J.B.: Two algorithms for automatic page layout and possible applications. Multimedia Tools Appl. 43(3), 275–301 (2009)
13. PicasaWeb, http://picasaweb.google.com
14. Platt, J.C.: AutoAlbum: clustering digital photographs using probabilistic model merging. In: Proceedings IEEE Workshop on Content-based Access of Image and Video Libraries, pp. 96–100 (2000)
15. Platt, J.C., Czerwinski, M., Field, B.A.: PhotoTOC: automatic clustering for browsing personal photographs. In: Proceedings of the 2003 Joint Conference of the Fourth International Conference on Information, Communications and Signal Processing, 2003 and the Fourth Pacific Rim Conference on Multimedia, vol. 1, pp. 6–10 (2003)
16. Rodden, K., Wood, K.R.: How do people manage their digital photographs? In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2003), pp. 409–416. ACM, New York (2003)
17. Sandhaus, P., Thieme, S., Boll, S.: Processes of photobook production. Special Issue of Multimedia Systems Journal on Canonical Processes of Media Production (2008)
18. Sarvas, R., Mäntylä, M., Turpeinen, M.: Human-Centric Design of Future Print Media. In: Proceedings of the PulPaper 2007 Conference, Helsinki, Finland (2007)
19. Snapfish, http://www.snapfish.com
20. Wang, W., Teo, P., Muzzolini, R.: Automatic digital photobook making system. In: Proceedings of the SPIE., vol. 7540, USA (2010)
21. West, D., Quigley, A., Kay, J.: MEMENTO: a digital-physical scrapbook for memory sharing. Personal Ubiquitous Computing 11(4), 313–328 (2007)
22. Xiao, J., Lyons, N., Atkins, C.B., Gao, Y., Chao, H., Zhang, X.: iPhotobook: creating photo books on mobile devices. In: Proceedings of the International Conference on Multimedia (MM 2010), pp. 1551–1554. ACM, New York (2010)

Icon Design for Older Users of Project Management Software

Christina Bröhl, Jennifer Bützler, Nicole Jochems, and Christopher M. Schlick

Institute of Industrial Engineering and Ergonomics, RWTH Aachen University
{c.broehl,j.buetzler,n.jochems,c.schlick}@iaw.rwth-aachen.de

Abstract. Working in projects is an important part of many jobs in service industry. Due to their knowledge and experience project planning is often accomplished by older employees. Therefore, and with regard to the demographic change an integration of the needs of older employees into the work environment is required. Common to most IT-based systems, including project management software, is the use of icons. To investigate different icon sets in project management software, regarding age related differences, two studies were conducted. The first study aimed at investigating two different icon sets regarding execution times and eye movements in an applied setting. The second study consisted of a questionnaire where subjects had to map different icons to their corresponding functions and had to compare these icons regarding their intuitiveness. Results revealed that older users profit from an icon design which is low in complexity but no impact by different icon designs was found for younger users.

Keywords: Usability, icon design, human-computer interaction, project management, eye-tracking, age differences.

1 Introduction

Well-organized project planning and subsequent execution of the project plan is nowadays a central predictor for the success of a project and thus a main competition factor for companies. Since projects are often linked to high risks as they are highly complex and constrained by tight time schedules [1], the planning process is often supported by computers. Therefore, the number of software products which support project planning is growing. These project management software applications offer different functions (e.g. project scheduling, cost planning, resource assignment planning, quality management, documentation) and differ in their visualization and planning methods (e.g. network diagram, Gantt chart). To navigate through the system and to generate different functions most commercial off-the-shelf project management software applications use icons. An icon is a small picture or a graphical symbol, which provides an illustration of a function or file on the system. Such an icon should transmit information about its underlying function in an abbreviated, simplified form and should not depend on letters or words [2]. Especially software tools which are

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 127–135, 2013.

unfamiliar to the user benefit from a well designed icon structure as the ease of use is facilitated which enhances the acceptance and further use of the tool.

Since project management places many demands on a person, like professional, methodological, expertise and life experience, it is often accomplished by older employees. Due to age-related changes, the interaction with computers is often particularly straining for older employees. These changes include decreases of fluid cognitive abilities like abstract reasoning [3], mental speed and short term memory storage [4], perceptual impairments like vision [5] as well as impairments in performing complex motor activities [6] which may lead to longer movement times [7] and difficulties in performing computer mouse tasks [8]. Although different people experience these impairments at different points in time and to different degrees, older people in general must deal with these functional limitations [9].

The aim of the current studies was an age differentiated usability evaluation of two different icon sets within a project management software application. The two icon sets that were used differed in their appearance: Icon set A included very simple icons which only consisted of three colors and were designed flat, whereas Icon set B included icons that had a more realistic appearance in that the icons were more colorful and had a glossy surface that was supposed to simulate three-dimensional objects. Within the first study (study I) participants generated typical project management tasks with both icon sets and results were analyzed regarding execution times and eye movements. The second study (study II) consisted of an online questionnaire where at first subjects had to map icons of both sets to their corresponding function and at second had to give a rating of icon intuitiviness.

2 Literature Review

2.1 Icon Design for Aging Users

Interfaces using icons aim at reducing complexity for the user as the mental workload is decreased when the icons are designed properly [10]. Icons represent information for the user graphically because they express semantic information not through the use of words but via nonverbal cues. Because of impairments in cognitive and perceptual functions which can go along with the aging process, older users may have difficulties in information processing [11] and hence in interpreting icons. Until now little has been reported in the research literature about icon design for the elderly. Mertens et al. [12] investigated the graphical presentation of icons and it could be shown that among elderly, the use of photos compared to pictograms or clip art leads to a higher recognition rate. However, the study was conducted using a hardcopy questionnaire and not a real user interface. McDougall, de Bruijn and Curry [13] found out that two parameters which play a critical role when studying icon design are the concreteness and the complexity of icons. Concreteness in this context means how closely the icon resembles items in the real world and complexity refers to amount of visual detail depicted in the icon. Schröder and Ziefle [11] investigated the effects of icon concreteness and complexity on semantic transparency in younger and

older users of mobile devices. Results showed that concreteness of an icon has a significant effect on subject's confirmatory response per item and that this effect does not vanish with familiarity of the icon. For icon complexity there was no significant difference in subject's confirmatory responses but subjects had higher reaction times when complex icons needed to be processed instead of simple ones. Regarding age related differences an effect for the processing time was found but no effect of the icon design so that the authors recommend a uniform design approach.

To build on existing research by exploring age-related differences in the effects of concreteness and complexity two different icon sets were investigated in this study which differed in the level of complexity and concreteness. Current research was extended by using icons of a commericial off-the-shelf project management software and studying the usability of these icons in an applied setting and in a questionnaire.

2.2 Eye-Tracking Metrics

Research has shown that eye-tracking measures can be used to detect usability problems in various contexts [14-16]. A number of eye movement metrics have been used to measure usability [16-17]. The mean fixation duration is one eye movement parameter that is often analyzed in usability studies. According to Goldberg and Kotval [14] longer fixation durations are an indicator for the difficulty to extract information from a display.

Concerning age differences in mean fixation durations most studies report higher mean fixation durations for older subjects. When analyzing navigational behavior on web pages, for example, Fukuda and Bubb [18] found that subjects aged between 62 and 74 years had longer fixation durations than younger participants aged between 17 and 29 years. Moreover, Hill et al. [19] investigated computer expertise when using the web within older subjects (70-93 years) and found that older novices had significant higher mean fixation durations than older experts.

Studies from other fields also address the effect of aging on eye-tracking metrics. In reading studies higher fixation durations were often found for older subjects [20-21]. However, there are also few studies that report no age differences regarding the fixation durations. Veiel et al. [22], for example, investigated age differences in the perception of visual stimuli and found no age difference regarding the fixation durations.

In line with these studies we used the mean fixation durations as dependent variable of the eye movements. Because the icons were arranged in the menu part of the software we furthermore analyzed the relative amount of fixation durations in menu parts.

3 Method

Dependent variables in study I were the execution time, the mean fixation durations and the relative amount of fixation durations in the menu area. Study II was analyzed

regarding errors made in a mapping task and regarding ratings in intuitiveness for the two different icon sets as dependent variables. Icons were selected from an existing commercial project management software and differed regarding their complexity and concreteness. Table 1 shows the icons of both sets used in this study. Icon set A consisted of very simply icons which were low in complexity and concreteness since the icons only consisted of three colors and were designed in bi-dimensional space. Icon set B included icons that had a more realistic appearance in that the icons were more colorful and had a glossy surface that was supposed to simulate three-dimensional space and was therefore rated as being high in complexity and concreteness.

Table 1. Icon sets investigated in the study

Meaning	Icon set A	Icon set B
Save		
Reload		
Undo		
Redo		
Settings		
Progress evaluation		
Insert activity		
Delete		
Upgrade		
Degrade		
Collective task		
Type of task		
Critical		
Done		
Started		
Relationship		
Zoom out		
Zoom in		

3.1 Participants

Altogether 10 subjects participated in study I and a sample of 19 subjects in study II. Subjects were divided in two age groups. In study I age ranged from 29 to 37 years in the younger age group (mean=32.00, SD=3.08) and from 43 to 60 in the older age group (mean=52.40, SD=7.13). In study II subjects in the younger age group were between 21 and 38 years old (mean=30.55, SD=4.80) and between 42 and 56 in the older age group (mean=49.25, SD=5.06). All participants had experience with project work and project management software but were not familiar with the design of the icons used in the study.

3.2 Apparatus

The experiment was conducted at a 22”-inch LCD TFT-widescreen-monitor. Eye movements were measured during the task using a Tobii X120 eye tracking system. Fixation durations were measured according to data typically found in the literature for usability studies, e.g. [16-17]. With regard to these studies the mimimum time for a fixation was set to be at least 100ms and the eye-in-head position had to be in a threshold of dispersion of about 2°.

3.3 Procedure and Task

Study I was conducted in a laboratory where subjects were seated in front of the monitor with a viewing distance of 500 mm. After calibrating the eye-tracking system, the subjects executed the tasks using the software with icon set A. This procedure was repeated for the software with icon set B as well. The representation of the icon set was permutated. The task consisted of typical steps (subtasks) when developing a project plan. Table 2 gives an overview of the nine sequential subtasks.

Table 2. Specification of the task

Subtask 1	Generate a new project plan
Subtask 2	Create five new activities and change duration of the activities
Subtask 3	Insert two new activities, change duration and level of the activities
Subtask 4	Insert four new activities, change duration and level of the activities
Subtask 5	Link activities
Subtask 6	Insert new milestone and link it to activity
Subtask 7	Insert new milestone and link it to activity
Subtask 8	Insert new milestone and link it to activity
Subtask 9	Assign responsible person

Study II was conducted using a questionnaire which could be completed online. In the first part of the questionnaire subjects had to map icons to their underlying function. Therefore, icons were arranged in blocks of 6 icons plus one icon which had no corresponding function serving as control item to make possible contradictions visible. In the second part of the questionnaire subjects had to rate icons according to their intuitiveness. Thereby, a function was given and subjects could choose between three answer options: Icon of set A, icon of set B and as a third option subjects could indicate whether the icon of set A or set B makes no difference to them.

3.4 Statistical Analysis

To analyze the data of study I statistically repeated measures analysis of variance was used. The age of the participants served as a between-group factor. Study II was analyzed using two-way ANOVA with age group and icon set as independent variables. Furthermore, a correlation analysis was used to investigate whether there is a relationship in the rating of intuitiveness and the performance in the mapping task. Because data for the correlation analysis were not measured at the interval level, correlation coefficients were calculated using Spearman's rho. According to Cohen & Cohen [23] effect sizes can be classified into small (r=.10), medium (r=.30) and high (r=0.50). The level of significance was set to $\alpha = 0.05$.

4 Results

4.1 Execution Times

To analyze the execution time similar subtasks were analyzed by combining the data. The analysis of the execution times resulted in significant main effect for the age group for subtasks 2-4 "Insert two new activities, change duration and level of the activities" ($F(1,8)=11,224$; $p=0,010$) and for subtasks 6-8 "Insert new milestone and link it to activity" ($F(1,8)=7,008$; $p=0,029$). In both tasks participants in the older age group needed significantly more time to execute the tasks than participants in the younger age group. For the main effect of the icon set as well as for the interaction effect between age group and icon set no significant effects were found.

4.2 Mean Fixation Durations and Relative Amount of Fixation Durations in Menu Parts

Mean fixation durations and the relative amount of the fixation durations in the menu part could not be analyzed with regard to age differences because data of the older age group were insufficient. This may be attributed to the problem of droopy eyelids. Droopy eyelids of participants when using eye-tracking metrics cause data loss because the eyelids or the eyelashes cover the pupil in specific gaze directions. The problem of droopy eyelids is a matter of individual differences and it grows with the age of participants [24]. Consequently, this might have been the case with data collected in this study.

Regarding the mean fixation durations no significant effect was found for the icon set when analyzing the graphical user interface as a whole. In order to investigate for which icon set more time was spent fixating it, the relative amount of fixation durations in the menu part where the icons are arranged was analyzed. Data revealed no significant effect when looking at the differences between the two icon sets when the mean fixation durations in the menu part were analyzed.

4.3 Questionnaire

The first part of the questionnaire was analyzed by measuring the performance in the mapping task for each icon set. Results indicate that subjects in the older age group made more errors with icon set B compared to subjects in the younger age group ($F(1,34)=8.816$; $p=0,005$). No age effect was found for icon set A.

When analyzing the second part of the questionnaire scores revealed that there was no significant difference in the intuitiveness ratings for one icon set or the other. To analyze whether intuitiveness ratings of the subjects are linked to the errors they made in the first part of the questionnaire a correlation analysis was conducted. According to the classification by Cohen and Cohen [23] a medium to high effect size was found for icons which were rated as being low in intuitiveness and errors made in the mapping task $r=.429$.

5 Discussion

The aim of the study was an age-differentiated analysis of two icon sets. The investigation of the execution times resulted in a significant main effect for the age group for two of the subtasks. In both tasks participants in the older age group needed significantly more time to execute the tasks than participants in the younger age group. This effect may be attributed to declines in perceptual and cognitive processes [3,4] like the hand eye coordination which may have an influence in this study because participants performed the tasks with a computer mouse. Regarding the execution time no effects were found for the icon set. So the complexity and the concreteness of the two investigated icon sets make no difference in accomplishing the subtasks used in this study. However, a tendency for the older age group was found in the scores to have shorter execution times when working with icon set A.

Results of eye-tracking data analyzed by means of fixation durations of the interface as a whole and the relative amount of fixation durations in menu parts did not differ for the two icon sets. As the fixation duration is assumed to be an indicator of the difficulty to extract information from a display it can be concluded that the design of the icons used in this study has no effect on the ability to extract information. Unfortunately, we were not able to investigate eye-tracking data with regard to age differences because we had difficulties in collecting data from older subjects. In future research we are trying to improve the measuring procedure so that an age differentiated analysis will be possible.

When analyzing the questionnaire results revealed that older subjects had more difficulties in the mapping task when working with icon set B. Together with the tendency for the older age group to have shorter execution times when working with icon set A this could lead to the conclusion that the older age group performs better with icons that are not high in complexity and concreteness. An explanation for this finding might be that the older age group already used to work with rather simple icons in the past since they grew up with early computers which used such an icon design in comparison to the younger age group, which showed no tendency for one icon set or the

other. Regarding the intuitiveness scores results showed that fewer errors were made in the mapping task with the icon set which was rated as being higher in intuitiveness. This finding is in accordance with existing literature [2] and shows that icon design places a crucial role in human-computer interaction.

Acknowledgements. The research was funded by the German Research Foundation according to the transfer project (SCHL 1805/6-1).

References

1. Shtub, A., Bard, J.F., Globerson, S.: Project Management: Processes, Methodologies, and Economics. Pearson Prentice Hall, Upper Saddle River (2005)
2. Horton, W.: Das Icon-Buch: Entwurf und Gestaltung visueller Symbole und Zeichen. Addison-Wesley, Bonn (1994)
3. Rodrigue, K.M., Kennedy, K.M.: The cognitive consequences of structural changes to the aging brain. In: Schaie, K.W., Willis, S.L. (eds.) Handbook of the Psychology of Aging, 7th edn., pp. 73–91. Elsevier, San Diego (2011)
4. Anstey, K.J., Low, L.F.: Normal cognitive changes in aging. Australian Family Physician 33, 783–787 (2004)
5. Goldberg, E.B.: Sensation & Perception. Thomson Wadsworth, Canada (2007)
6. Newell, K.M., Vaillancourt, D.E., Sosnoff, J.J.: Aging, complexity, and motor performance. In: Birren, J.E., Schaie, K.W. (eds.) Handbook of the Psychology of Aging, 6th edn., pp. 163–182. Elsevier, San Diego (2006)
7. Cerella, J.: Aging and information processing rate. In: Birren, J.E., Schaie, K.W. (eds.) Handbook of the Psychology of Aging, pp. 201–221. Academic Press, New York (1990)
8. Smith, M.W., Sharit, J., Czaja, S.J.: Aging, motor control, and the performance of computer mouse tasks. Human Factors 41, 389–396 (1999)
9. Pijukkana, K., Sahachaisaeree, N.: Graphical Design and Functional Perception on Technology-Driven Products: Case Study on Mobile Usage of the Elderly. Procedia - Social and Behavioral Sciences 42, 264–270 (2010)
10. Lodding, K.: Iconic interfacing. Graph. Appl. 3, 11–20 (1983)
11. Schröder, S., Ziefle, M.: Effects of icon concreteness and complexity on semantic transparency: Younger vs. Older users. In: Miesenberger, K., Klaus, J., Zagler, W.L., Karshmer, A.I. (eds.) ICCHP 2008. LNCS, vol. 5105, pp. 90–97. Springer, Heidelberg (2008)
12. Mertens, A., Brandl, C., Przybysz, P., Koch-Körfges, D., Schlick, C.: Design recommendations for the creation of icons for the elderly. In: Soares, M., Jacobs, K. (eds.) Work: A Journal of Prevention, Assessment and Rehabilitation - IEA 2012: 18th World congress on Ergonomics - Designing a Sustainable Future, Amsterdam, p. 41 (2012)
13. McDougall, S.J., de Bruijn, O., Curry, M.B.: Exploring the effects of icon characteristics on user performance: the role of icon concreteness, complexity, and distinctiveness. Journal of Experimental Psychology Applied 6(4), 291–306 (2000)
14. Goldberg, J., Kotval, X.P.: Computer interface evaluation using eye movements: methods and constructs. International Journal of Industrial Ergonomics 24(6), 631–645 (1999)
15. Cowen, L., Ball, L.J., Delin, J.: An eye movement analysis of webpage usability. In: Faulkner, X., Finlay, J., Detienne, F. (eds.) People and Computers XVI - Memorable yet Invisible, pp. 317–335. Springer, London (2002)

16. Jacob, R.J.K., Karn, K.S.: Eye tracking in human–computer interaction and usability research: Ready to deliver the promises. In: Hyona, J., Radach, R., Deubel, H. (eds.) The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research, pp. 573–605. Elsevier (2003)
17. Poole, A., Ball, L.J.: Eye tracking in HCI and usability research: Current status and future prospects. In: Gahoui, C. (ed.) Encyclopedia of Human-Computer-Interaction, pp. 211–219. Idea Group Reference, Hersey (2005)
18. Fukuda, R., Bubb, H.: Eye tracking study on web-use: Comparison between younger and elderly users in case of search task with electronic timetable service. PsychNology Journal 1(3), 202–228 (2003)
19. Hill, R., Dickinson, A., Arnott, J., Gregor, P., McIver, L.: Older web users' eye movements: experience counts. In: Proc of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1151–1160. ACM, Vancouver (2011)
20. Kliegl, R., Grabner, E., Rolfs, M., Engbert, R.: Length, frequency, and predictability effects of words on eye movements in reading. European Journal of Cognitive Psychology 16(1-2), 262–284 (2004)
21. Rayner, K., Castelhano, M., Yang, J.: Preview benefit during eye fixations in reading for older and younger readers. Psychology and Aging 25(3), 714–718 (2010)
22. Veiel, L.L., Storandt, M., Abrams, R.A.: Visual search for change in older adults. Psychology and Aging 21(4), 754–762 (2006)
23. Cohen, J., Cohen, P.: Applied multiple regression/correlation for the behavioral sciences. Erlbaum, Hillsdale (1983)
24. Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., van de Weijer, J.: Eye tracking: A comprehensive guide to methods and measures. Oxford University Press, Oxford (2011)

Search Strategies in Hypermedia Navigation and Spatial Abilities: A Comparison with Physical Navigation

Aurélie Brouwers

Research Fellow F.R.S-F.N.R.S, Université catholique de Louvain, Belgium
aurelie.brouwers@uclouvain.be

Abstract. This article focuses on spatial abilities mobilized during hypertext navigation. Based on the evidence that spatial cognition plays a central role in navigation, we present an experiment involving information search tasks both in physical environment and in hypermedia environment. We investigate how users make use of their spatial abilities to search information in hypermedia, by comparing their performances in hypermedia navigation and physical navigation. As data collection and analysis are still in progress, we present preliminary results based on available data.

Keywords: hypermedia, navigation, visuospatial abilities, spatial cognition.

1 Introduction

The term "navigation" has been used for decades to describe the interaction between a user and a hypertext system [1]. The use of this spatial metaphor is backed by at least two kinds of research results. On the one hand, a number of studies [2]–[4] have uncovered correlations between subject performance in hypertext navigation and performance in standardized tests of visuospatial cognitive abilities. On the other hand, hypertext users have been shown to use spatial metaphors extensively and systematically when they speak of their interaction with hypertext [5], [6], indicating that these metaphors play a role in the way users think about this interaction.

This article focuses on the role of spatial cognition and specifically visuospatial abilities in information search in a hypertext environment. We investigate how users make use of their spatial abilities to search information in hypermedia, by comparing their performances in hypermedia navigation and physical navigation. First, we will review relevant research in spatial cognition and information search. Next, we will present an experiment requiring subjects to interact with a physical environment and a hypertext environment. As data collection and analysis are still in progress, we will present preliminary results based on available date.

2 Spatial Cognition and Cognitive Mapping

The concept of cognitive map is central to the field of spatial cognition research. Siegel and White distinguished between landmarks, routes and survey representations as

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 136–145, 2013.

components of cognitive maps. Landmarks are features of the environment that are saillant to the subject. Routes connect different landmarks. Survey knowledge organizes routes into configurations that provide the subject with an overview of their environment. Siegel and White described a "Main Sequence" for the development of spatial knowledge that goes "from landmarks, to route-maps, to survey-maps" as "a process of going from association to structure, and of deriving simultaneity from successively" [7]. As Siegel and White [7] pointed, cognitive maps are not very much "map-like": they tend to be fragmented and are prone to distortion. Hence, we prefer to speak about spatial mental representations.

Moeser [8] compared the navigation of student nurses who has been working in hospital for three years and the navigation of naive subjects who studied the hospital map. He showed that the naive subjects performed significantly better on objective measures of cognitive mapping (distance estimation, drawing plan, etc.) than did the nurses. He concluded that the extended and intensive experience of an environment doesn't systematically provide the individual with survey knowledge of it. Each person tends to construct mental representations that include relevant elements for their own use of the environment. For these student nurses, landmarks and route knowledge seemed sufficient to navigate their professional environment.

In hypermedia research, the concept of cognitive map has been used to describe the user's mental representation of the hyperdocument's organization [9]–[12]. Following Siegel and White's model, several studies in hypermedia research (e.g. Sedig et al.[13]; Kim & Hirtle [11]; Edwards & Hardman [14]) relied on the view that landmarks, routes and survey knowledge are acquired successively. This implies that survey knowledge is necessarily the most advanced form of spatial knowledge. Kim & Hirtle [11] and Edwards and Hardman [14] emphasized that subjects with survey knowledge are rarely disoriented in hypermedia.

Spatial mental representations of the environment can be constructed either through primary learning (i.e. by observing the actual environment directly) or through secondary learning (i.e. by means of an external presentation, such as a map) [15]. Secondary learning may represent an alternative way of acquiring spatial knowledge (compared to Siegel and White's "Main Sequence"), as survey knowledge may be acquired directly from a map, without being derived from route knowledge acquired through navigation. While this distinction is relevant to physical environments, its application to hypermedia environments is problematic, as the organization of hyperdocuments only exists in representations, be they external (e.g. concept maps, or the system's interface itself) or internal (the mental representation constructed by the user). Hence, the very notion of primary learning seems void in the context of hypermedia. However, the notion of learning routes from navigating (which can in turn be elaborated into survey knowledge) as opposed to learning survey configurations from a site map (which in turn can be converted into specific routes from one page to another), stays relevant in this context.

People differ in their spatial abilities. Research by Goldin and Thordyke [16] showed that the difference between "poor cognitive mappers" (subjects with lower spatial abilities) and "good cognitive mappers" (subjects with higher spatial abilities)

may lie in their ability to construct spatial mental representations of their environment, both from navigation and from maps. However, when forced to study a map so as to construct a reliable representation of a given environment, "poor cognitive mappers" navigated as efficiently as "good cognitive mappers" suggesting that "when utilizing equally accurate knowledge, poor cognitive mappers can navigate as well as good cognitive mappers"[16].

Spatial visualization abilities of hypermedia users had been shown to predict their navigation performance by previous studies [2]–[4]. Spatial visualization can be defined as an internal spatial ability. Such internal abilities, requiring "a purely mental effort to obtain the correct answer", were distinguished by Dahlbäck et al. [4] from external spatial abilities that involve the active manipulation of physical objects (such as the arrangement of block in a pattern defined by a picture of arranged blocks), which did not correlate with performance in hypertext information search.

Carroll [17] showed that the spatial visualization factor assessed by these tests involves the cognitive processing required to mentally encode and manipulate spatial shapes. He also noted that the successful completion of tests such as VZ-1 and VZ-2 [18] requires a mental representation in three dimensions. This would explain his observation that individuals who perform well in task in 2-D environments tend to be also performing well in tasks in 3-D environments.

Interestingly, the tests used to assess spatial visualization abilities actually require subjects to process external representations (pictures) into mental representations they need to manipulate internally in order to complete the task. Constructing a reliable mental model from an external representation (as opposed to using a mental model to act in the real world, or to using an external representation to act in the real world).

Research on the use of concept maps to represent hypermedia structure has yielded ambiguous results. Dee-Lucas & Larkin [19] showed that a structured overview of hypertext had beneficial effects on the memorization on the document's textual contents. Vörös et al. [20] showed that concept maps helped subjects with poor spatial abilities to better remember the hypermedia structure after their navigation. However, their study did not test whether subjects with more accurate cognitive maps of the hypertext navigated more efficiently (i.e. performed better on search tasks). In Nilsson & Mayer's [21] studies, subjects who used a hyperdocument with a non-clickable map performed better than subjects who used the same hyperdocument without the map at first, but the comparison inverted over time, showing the map could become cumbersome during navigation. Scott & Schwartz [22] showed that while processing a hypertext concept map generated additional cognitive load, the extra effort paid off when the spatial relationships depicted by the map matched the semantic relationship of the hypertext's contents.

3 Research Question and Hypothesis

Our research question is the following how do spatial visualization abilities come to play in hypermedia navigation?

Spatial visualization abilities could play a role (1) in the construction of a mental model of the hypertext's organization, (2) in the use of this mental model to make navigational choices, or (3) in the use of the hypertext interface and navigation tools (as an external representation of its organization) to make navigational choices.

Based on our literature review, we hypothesize that spatial visualization abilities mainly play a role during the construction of the hypertext mental model. To test this hypothesis, we set up an experiment requiring subjects to perform a series of information search tasks in a physical and a hypertext environment. Our experimental protocol allows us to proceed with intra-subject comparisons of navigation behaviors in hypertext and in real life, as well as inter-subjects comparisons based on their level of spatial visualization abilities (cf. infra). It also allows us to clarify the role of map usage during hypertext navigation.

4 Method

4.1 Fields of Observation

Our experimental protocol includes tasks to be completed in a large-scale physical environment and a hypermedia environment. The Planckendael (Belgium) zoo was chosen as the physical environment, as it met all of our requirements. In addition to making different animals available for public display, the zoo includes and organizes information (about animals) spatially (e.g. in the form of posters presented at each animal's enclosure). It clearly bounded space corresponds to a bounded informational space. In this context, information search tasks can be designed, involving the processing of both spatial and semantic information by subjects. Also, it is an outdoor environment, which allows GPS tracking of our subjects.

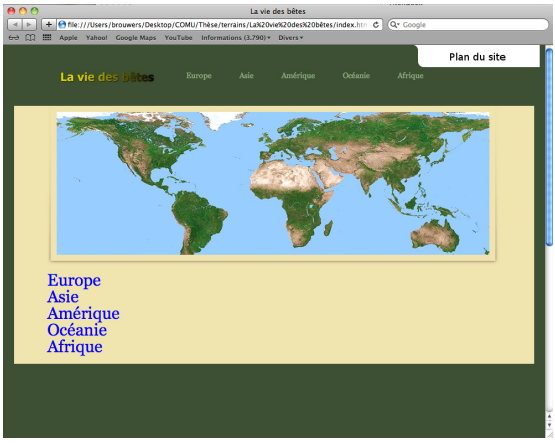


Fig. 1. Home page

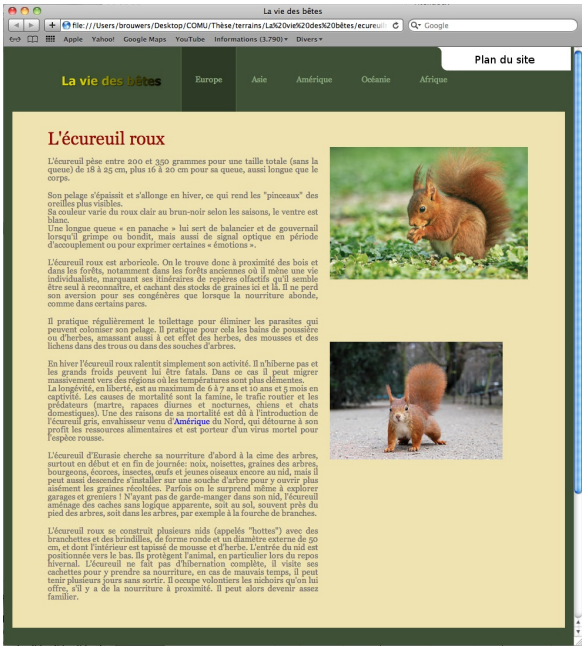


Fig. 2. Animal page

We designed a forty-five-page hypermedia environment on wildlife, including information on different animal species that are not visible at the zoo. Each animal is presented on a separate page. Pages presenting an animal species are grouped according to the continent on which the species lives. This hierarchical structure mimics that of the zoo, which dedicates a part of its space to each continent.

4.2 Participants

Students enrolled at the Université Catholique de Louvain (Belgium) are currently being recruited as voluntary participants in our experiment. Ten subjects have completed the experiment so far. Twenty more subjects will be tested in the near future.

4.3 Individual Variables

Subjects are tested with respect to two types of cognitive abilities. On the one hand, we use VZ-1 and VZ-2 tests from the kit of factor-referenced cognitive tests [18] to assess our subjects spatial visualization abilities.

On the other hand, we test our subjects with a discourse comprehension test [23]. As information search in hypermedia involves the processing of both semantic and spatial information [24], we use this test to control our subjects verbal abilities.

4.4 Tasks

Subjects are required to perform five information search tasks in the physical environment, and seven information search tasks in the hypertext environment. For each task subjects are instructed to answer a specific question about an animal species, the answer of which is located at the animal's enclosure in the zoo, or on the animal's page in the hypertext. The questions used in the two environments are different.

4.5 Procedure

Prior to the navigation, half of the subjects (randomly selected) are asked to study the map of the environment they are about to interact with (subjects either study both maps, or none). Subjects are tested on their memorized map, and the study is repeated until they learned the map perfectly. According to our hypothesis, subjects with low visual-spatial abilities who studied the map should perform similarly to with high visual-spatial abilities, as visuospatial abilities are supposed to play a decisive role in the construction of the mental model of the environment.

During the navigation, subjects are provided with the task questions one at a time. When the subject thinks to have found the answer, he notifies the experimenter. If the answer is correct, he is given the second task question. If not, the subject has to continue searching. While searching subjects are asked to verbalize their navigation intentions before executing them. In physical environment, subjects are asked to estimate the direction of the searched information. During navigation, subjects have the opportunity to consult a map. The hypertext map can be accessed though a tab (preventing concurrent navigation and map viewing) and in the zoo, subjects are required to stop whenever they want to view the map. The aim is to force subjects to verbalize their use of the map.

In each condition (map-study vs. no-map-study), half the subjects navigate the zoo prior to the hypertext, and the other half navigate the hypertext prior to the zoo, in order to control the training effect of one navigation on the other.

Finally, subjects are submitted to the spatial and verbal ability tests after the two navigation sessions, in order to avoid a sense of demotivation due to poor performance in these tests.

4.6 Data Collection

In the zoo, navigation is recorded using a Looxcie, a mini camera equipped with a microphone. A GPS tracker, Map My run, an android application, records the routes. Hypermedia navigation tasks are recorded using TechSmith Morae, a software suite that captures screen and webcam activity, as well as mouse and keyboard input (e.g. clicks, menu selection, etc.).

5 Preliminary Results

At the time of writing, we have tested ten subjects and have started to analyze the collected data for nine of them. Our current analysis focuses on the influence of the study of the map on navigation behavior and efficiency. Table 1 presents some of the data that were collected for these subjects.

Table 1. First data

			Zoo			Hypertext	
Sub-ject	Condition	VZ-1+ VZ-2 score	Total search time	Total distance	Map views	Total search time	Map views
1	no-map-study	1	01:08:00	5,57 km	7	00:28:38	0
2	no-map-study	6	00:57:35	3,59 km	12	00:11:40	10
3	no-map-study	12	00:45:00	6,49 km	8	00:27:57	0
4	no-map-study	16	00:59:21	4,60 km	6	00:09:58	0
5	no-map-study	17	01:12:00	4,13 km	19	00:30:00	0
6	map-study	2	01:10:00	5,50 km	12	00:12:18	0
7	map-study	12	00:38:42	2,54 km	4	00:11:30	7
8	map-study	19	01:04:00	6,68 km	8	00:09:09	0
9	map-study	22	00:21:08	3,54 km	7	00:18:57	0

Contrary to previous research results, scores to the spatial visualization tests did not correlate significantly with time spent navigating in either of the environments, or with distance traveled in the zoo (using Pearson's correlation coefficient), although this may be due to the low number of subjects (N=9) in our current dataset.

Our current observations for the physical environment are not consistent with previous studies [16]. Indeed, subjects who studied the map of the zoo are not necessarily more efficient than others ($Mean_{map\text{-}study}$ = 48:28; $Mean_{no\text{-}map\text{-}study}$ = 1:00:23). However, subjects who studied the site map prior to navigation completed their tasks more quickly than the other subjects ($Mean_{map\text{-}study}$ = 12:58; $Mean_{no\text{-}map\text{-}study}$ = 21:39).

We observed that the subject with the lowest results in VZ-1 and VZ-2 test in the map-study condition was able to use her survey knowledge in the zoo, as she was able to accurately estimate the direction of the searched item. However, she was unable to convert her survey knowledge into a proper route plan, as the geodesic distance between her current location and her planned destination did not correspond to an actual path in the zoo. In this case, a survey vision of the environment proved to be of little help to our subject. In the hypertext environment, she successively browsed the main nodes of the interface (the continent pages) to gain visual access to the different links (routes) to specific animal pages, until she identified a potential target node. In this case, she used her memory of the overall (survey) organizing principle of the

hypertext to structure her use of the interface (as an external representation of the system's structure).

Finally, the majority of our subjects did not use the sitemap tab while navigating. When asked why they didn't, most subjects replied that they feared the concept map would confuse them, as they felt sitemaps are generally useless.

6 Conclusions and Future Work

The experiment presented in this paper aims at better understanding the role of visuospatial abilities in hypertext navigation. To do this, we compare the information search strategies developed by subjects in a physical environment and in a hypertext.

Observation and analysis are still in progress. Nevertheless, preliminary results indicate that the study of the map before hypertext navigation does make navigation easier for individuals with low visual-spatial abilities. This suggests that the visuospatial abilities may play a more decisive role in the construction of the mental representation of the environment than in the use of this representation during navigation.

Future qualitative analyses of our data will attempt to identify and compare strategies for navigating our two environments. Specifically, we intend to identify (1) the type of information subjects use to make navigation choices (route vs. survey knowledge; their mental model vs. the map vs. cues in the environment) and (2) the cognitive processes they perform during navigation to use this information, e.g. coordinating the map with the territory, or converting survey knowledge into a route plan. As far as the effect of spatial abilities is concerned, we intend to observe whether subjects with low spatial abilities develop only route knowledge, whether studying the map prior to navigation allows subjects with low spatial abilities to develop survey knowledge, and whether they are able to use it to navigate the environment.

As part of our analyses, we will diversify the indicators we use to assess navigation efficiency. We will use the efficiency metric proposed by Smith (1996) for hypertext navigation, which combines three kinds of indicators: a measure of the redundancy (repeated visits to the same nodes) of navigation, a ratio between the number of nodes required to complete a task and the number of different nodes visited, and an indicator of the successful completion of the task. We are working on an adaptation of this metric to the physical environment, using the traveled distance instead of the number of visited hypertext nodes.

We are hopeful that our work will yield recommendations for hypertext designers on how to provide users with low spatial abilities with survey information they can easily convert into route plans.

References

1. Parunak, H.V.D.: Hypermedia topologies and user navigation. In: HYPERTEXT 1989, Proceedings of the Second Annual ACM Conference on Hypertext, pp. 43–50. ACM, New-York (1989)

2. Blustein, J., Ahmed, I., Parvaiz, H., Fu, C.-L., Wang, C., Hu, Y.: Impact of spatial visualization aptitude on WWW navigation. In: Proceedings of the Workshop on Cognition and the Web 2008, Granada, pp. 185–192 (2008)
3. Campagnoni, F.R., Ehrlich, K.: Information retrieval using a hypertext-based help system. ACM Transactions on Information Systems 7(3), 271–291 (1989)
4. Dahlbäck, N., Höök, K., Sjölinder, M.: Spatial Cognition in the Mind and in the World the Case of Hypermedia Navigation. In: Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society, Mahwah, pp. 195–200 (1996)
5. Maglio, P.P., Matlock, T.: The Conceptual Structure of Information Space. In: Höök, K., Benyon, D., Munro, A.J. (eds.) Designing Information Spaces: The Social Navigation Approach, pp. 385–404. Springer, New-York (2003)
6. Collard, A.-S., Fastrez, P.: A model of the role of conceptual metaphors in hypermedia comprehension. In: Proceedings of CICOM: Communication Sciences International Congress on Communication, Cognition and Media, pp. 241–255 (2009)
7. Siegel, A.W., White, S.: The development of spatial representations of large-scale environments. In: Advances in Child Development and Behavior. Academic Press, New York (1975)
8. Moeser, S.D.: Cognitive Mapping in a Complex Building. Environment and Behavior 20(1), 21–49 (1988)
9. Dillon, A., McKnight, C., Richardson, J.: Navigation in hypertext: a critical review of the concept, Amsterdam, New York, pp. 587–592 (1990)
10. Dillon, A., McKnight, C., Richardson, J.: Space - The final chapter or Why physical representations are not semantic intentions. In: McKnight, C., Dillon, A., Richardson, J. (eds.) Hypertext. A Psychological Perspective, pp. 169–191. Ellis Horwood Ltd. (1993)
11. Kim, H., Hirtle, S.C.: Spatial metaphors and disorientation in hypertext browsing. Behaviour & Information Technology 14, 239–250 (1995)
12. Shum, S.: Real and Virtual Spaces: Mapping from Spatial Cognition to Hypertext. Hypermedia 2(2), 133–158 (1990)
13. Sedig, K., Rowhani, S., Liang, H.-N.: Designing interfaces that support formation of cognitive maps of transitional processes: an empirical study. Interacting with Computers 17(4), 419–452 (2005)
14. Edwards, D.M., Hardman, L.: Lost in Hyperspace". Cognitive mapping and navigation in a hypertext environment. In: McAleese, R. (ed.) Hypertext: Theory into Practice, ch. 7, pp. 105–125. Blackwell Scientific Publications Ltd., Oxford (1989)
15. McDonald, T.P., Pellegrino, J.W.: Psychological perspectives on spatial cognition. In: Gärling, T., Golledge, R.G. (eds.) Behavior and Environment. Psychological and Geographical Approaches, pp. 47–82. North Holland, Amsterdam (1993)
16. Goldin, S.E., Thorndyke, P.W.: Spatial Learning and reasoning skill. Rand Corporation (1983)
17. Carroll, J.B.: Human cognitive abilities: a survey of factor-analytic studies. Cambridge University Press (1993)
18. Ekstrom, R.B., French, J.W., Harmon, H.H.: Kit of Factor-referenced Cognitive Tests. Educational Testing Service, Princeton (1976)
19. Dee-Lucas, D.: Effects of overview structure on study strategies and text representations for instructional hypertext. In: Rouet, J.-F., Levonen, J.J., Dillon, A., Spiro, R.J. (eds.) Hypertext and Cognition, pp. 73–108. Lawrence Erlbaum Associates Inc., Mahwah (1996)
20. Vörös, Z., Rouet, J.-F., Plèh, C.: Content maps help low spatial capacity users memorize link structures in hypertext. In: Proceedings of the Workshop on Cognition and the Web 2008, Granada, pp. 185–192 (2008)

21. Nilsson, R.M., Mayer, R.E.: The effects of graphic organizers giving cues to the structure of a hypertext document on users navigation strategies and performance. International Journal of Human-Computer Studies 57, 1–26 (2002)
22. Scott, B.M., Schwartz, N.H.: Navigational spatial displays: The role of metacognition as cognitive load. Learning and Instruction 17(1), 89–105 (2007)
23. Boutard, C., Claire, I., Gretchanovsky, L.: Le vol du P.C.: évaluation fonctionnelle de la lecture chez les jeunes de 11 à 18 ans. Ortho edn. (2006)
24. Dillon, A.: Spatial-Semantics. How users derive shape from information space. Journal of the American Society for Information Science 51, 521–528 (2000)

Data Visualisation, User Experience and Context: A Case Study from Fantasy Sport

Rob Euman and José Abdelnour-Nocera

Sociotechnical Centre for Internationalisation and User Experience
University of West London, Ealing, W55RF, 69121 United Kingdom
rob.euman@lab49.com, jose.abdelnour-nocera@uwl.ac.uk

Abstract. Fantasy Football is a rapidly growing online social game. As users become more sophisticated and technology advances, the amount of data that is available to inform users' decision making is growing rapidly. Representing this data in an informative and engaging way can be a challenge but data visualisation offers many ways to achieve this. This paper focuses on the design of interactive solutions that are measured against existing products by way of a comparative evaluation. In order to study the impact on user performance, efficiency and accuracy are measured for clearly defined tasks carried out on each design. The user experience is measured to understand the satisfaction and perceived ease of use of each visualisation system. This study will be useful to validate or challenge existing principles of data visualisation design and perception as well as offering suggestions for improving fantasy football products. The study will also serve as a case study to support further research into data visualisation evaluation methods. The paper concludes by discussing the findings and possible areas for further research and design.

Keywords: HCI, social games, data visualisation, design, evaluation.

1 Introduction

Fantasy sports users are becoming increasingly sophisticated in the way they consume and analyse sports data. However, the ways in which this data is represented has not evolved as much as in other fields. There may be opportunities for the fantasy sports world to learn from domains such as finance, science and healthcare by using data visualisation techniques to improve the performance and experience of its users.

The design rationale for this study follows the hypothesis that data visualisation can improve the both the performance and experience of fantasy football users. In challenging this assumption, empirical research was conducted involving real fantasy football users to understand their current experience and ways in which this could be improved. The research was conducted over a limited time period and focused on the English Premier League season of 2011/12.

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 146–155, 2013.

A comparative study was conducted which required a solid understanding of existing fantasy football games that fairly represents the existing fantasy football experience and comparing this with new systems that were designed and prototyped as part of the research process. Although the prototypes evaluated represent only a subset of a fully functioning product, they were adequate as concepts and included enough functionality to test and challenge the hypothesis.

2 Existing Literature

2.1 Data Visualisation

Information Visualisation is the accepted term for the field of graphical communication that data visualisation sits within. Underpinning the design solutions and data analysis in this study are well established principles of human perception and statistical graphics. Clevelend & McGill [1] wrote a foundational paper in this field where visual elements such as colour, size and shape were tested to measure their effectiveness in communicating quantitative information. In addition to this, Bertin's *Semiology of Graphics* [2] looks at the language of visualisation techniques and Ware's *Information Visualisation* [3] looks at the subject from the perspective of perceptual science. Almost all research on this subject will reference the work of Edward Tufte who's books [4, 5] are ground breaking in their approach to information presentation and visual literacy.

The field of interactive data visualisation is broad and continuously evolving due to advances in technology. Ben Shneiderman has been responsible for many milestones in this research field. The paper *Visual Information Seeking* [6] is often referenced as one of the first papers to consider visualisations beyond static representations of data and promoted the benefits of interactivity. In the paper *The Eyes Have It,* Shneiderman [7] looks at a taxonomy system for interactive visualisation techniques that are categorised by data type.

2.2 Data Visualisation Evaluation

As demand for new ways of visualising data increases, so does the need for standards and guidelines to support their creation. There is a growing body of work that focuses purely on data visualisation evaluation and the challenges this field faces that are not common in standard interactive systems evaluation [8]. Plaisant [9] addresses these challenges and recommends improving access to repositories of data, toolkits and development tools and case studies that could encourage greater investment in the field. Carpendale [10] discusses the trade-off between generalisability, precision and realism and recommends a mixed methods approach to evaluations which includes both qualitative and quantitative research. Lam et al [11] recommend evaluation approaches specific to seven specific scenarios of which *Evaluating User Performance* and *Evaluating User Experience* are key to this study.

3 Methodology

3.1 Evaluation Approach

The main challenge in evaluating data visualisation systems lies in the fact that it is difficult to replicate the conditions in which these systems are used. A standard usability evaluation may involve a controlled or informal study of a user interacting with a particular system. Tasks are often brief and satisfaction, efficiency and effectiveness can all be measured within the test. Systems that involve large amounts of data, whether visualised or not, are usually used in a more exploratory way and discovery is a key part of the experience. It may take some time before the user completes a task such as finding a specific piece of information amongst the data and this can happen long after use of the visualisation system. Furthermore, a visualisation system's effectiveness may lie in answering questions you didn't know you had [9].

In order to avoid some of these common pitfalls of data visualisation evaluation, this research has used a combination of methods that are supported by previous research and case studies. As there is no universally applicable methodology for evaluating data visualisation systems, care has been taken to focus on scenarios that are relevant to the context of this study. This study focuses on two scenarios in challenging the above hypothesis. These are *Evaluating User Performance* and *Evaluating User Experience* [13].

3.2 Research Design

In this study three ways of visualising fantasy football data in an interactive context are compared in the form of interactive prototypes. These prototypes represent only a subset of a fantasy football game as the study only focuses on the visual representation of player data. Taking this approach is also beneficial as a fully horizontal prototype containing unnecessary functionality would distract the user and take up unnecessary design and development time.

Prototype A (Fig. 1) is a table that reflects the conventional format of existing fantasy football products. This prototype served as a benchmark to measure the other two data visualisation techniques against. Prototype B uses the same table format but with the player points data from each match of the season represented in a graphical format. Prototype C uses the same dataset as A and B but it is presented as a treemap. The comparison involved carrying out a simultaneous study of each interface using real fantasy football users carrying out tasks reflective of their normal pattern of use.

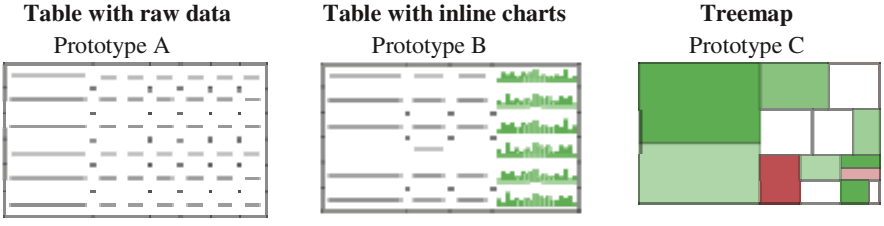


Fig. 1. Three prototypes to be evaluated

3.3 Identification of Variables

To maximise the accuracy of this evaluation, each of the variables has been identified (Table 1). Data visualisation techniques, represented by the three prototypes, are the independent variables that have been managed throughout the evaluation. The dependant variables of user performance and experience were measured in relation to the independent variable. There were a number of controlled variables that did not change throughout the evaluation. Where it was not possible to control these variables, it is acknowledged and considered in the analysis of the results.

Table 1. Independent, Dependant and Controlled Variables

Independent	Dependant	Controlled
Data visualisation technique: Table with raw data, table with inline charts, treemap.	**User Performance:** Task completion time, task error rates. **User experience:** Perceived ease of use, satisfaction.	**Participants:** Domain knowledge, technical ability. **Evaluation:** Tasks, questions, device, environment.

3.4 Evaluation Methods

In order to make valid judgments on the three prototypes they were measured against specific metrics (Table 2) through a variety of evaluation methods. The mixed methods approach helped capture both objective and subjective information in order to measure the impact each visualisation technique has on user performance and experience.

Table 2. Metrics to be Measured Throughout the Evaluation

User Performance	User Experience
Completion time: The efficiency of the system as participants complete the tasks.	**Ease of use:** The participants' perceived ease of use.
Error rates: The effectiveness and tolerance of the system.	**Satisfaction:** Whether the system was enjoyable to use and likelihood of adoption.

Quantitative Methods

Timed Tasks. These are scripted tasks that ask participants to interact with each prototype in a controlled way. Each task reflected that of a real session such as making selections and comparisons based on fantasy player points and prices. The tasks were ordered in such a way that avoided bias toward either prototype or participant.

Error Rates. As part of the controlled experiment, the amount of tasks that are failed or abandoned was also measured. Errors are highlighted as significant issues when measuring the performance of the system.

Qualitative Methods

Observations. Behaviours were observed as the participants performed each task. These are brief, subjective and unobtrusive with the participant being asked to focus on completing the task instead of discussing their actions at the same time.

Voice Mapping. This is a technique for measuring participants' attitudes towards the system. It is particularly useful when making comparisons between systems or measuring the effect of design improvements over time. The participants were asked to plot a mark on an axis (measuring satisfaction and perceived ease of use.

3.5 Prototype Design

Tables. Although the table format reflects existing fantasy football interfaces, there are still a number of variables that need to be controlled such as sorting, filtering and visual design. There are 531 rows, one for each Premier League player. Columns contain alphabetic data including player name, club and position as well as numerical data such as total player points, average player points, player value and games played. The only difference between the two tables is the representation of player points per match. Prototype A lists points as numerical values in a column. Prototype B represents this data both as numerical data and as a bar chart representing each of the 38 matches of the football season (although few players actually made this many appearances).

531 rows of data would have a negative impact on cognitive and page load so it is important for the user to be able to control the amount and order of this information. The default view shows 10 players while a dropdown allows the user to change this to 50 or 100 players. The column headers also allow the user to sort the data in the table. Columns containing individual match player points have had this functionality removed as users are only interested in the cumulative player point totals. However, position, team, player value and points are all categories that users like to sort by.

There were opportunities to improve on the presentation of these tables by applying established design principles of information design. However, this would introduce too many variables to the evaluation and distance both prototype A and B from the existing fantasy football product table design.

Treemaps. Treemaps use colour, size and proximity to represent relevant attributes of the data. This is represented in the form of a 2 dimensional display of nested rectangles. Within each rectangle is a collection of tiled rectangles with varying sizes and colours. The information represented by the colour and size of these rectangles can vary depending on domain and dataset. For the purpose of this study I have chosen to focus on the typical financial use of treemaps where the size of the rectangle represents volume or value while the colour of the rectangle represents percentage change or stock performance.

The Squarified treemap [12] format lends itself well to compare, interact with and perceive structure. While the Squarified algorithm is limited in its ability to order the

rectangles, This approach was chosen to create the Fantasy Football treemap as the order of player values is less relevant to the size and colour of the rectangles.

Ultimately fantasy football users are looking for information on the performance of individual players. However, the routes users take to access this information can differ. Some users are looking for the best value players in a specific position, others are looking for the most consistent players of a specific team and others may simply be looking for a specific player. Therefore, it is important to use the hierarchy to allow users to make sense of the data format and access the information they need.

The player performance data in the fantasy football treemap is represented by a graduation of colour from green for positive values to red for negative values. To provide contrast and to avoid undesirable colour blends, white is used as a midpoint. Colours are displayed not on actual points scored but on each player's points in relation to the average point score. This has been done as most players' scores are a positive amount (i.e., the highest score is 2451 and the lowest is -15) which means the treemap will display predominantly green rectangles which will only use around 50% of the colour scale. This will make comparisons more difficult. By assigning colours for player points in relation to the average player point figure there should be an even split of red and green rectangles.

While colour is an effective technique for conveying information, around 10% of the male population suffers from some form of colour blindness. For this reason, an option to change the treemap to use the colours yellow and blue to represent player points should be included in a final product.

3.6 Evaluation

The tasks used in the evaluation were chosen carefully to insure accurate and consistent data could be captured and to reflect the common tasks of the participants' normal usage. It was also important that each participant did not carry out the same task on more than one prototype. This would skew the results as, once a participant had found the information via one prototype, the same task on another prototype would be one of navigation instead of discovery.

Nine tasks were chosen which could be categorised into three groups; overview level tasks, team level tasks and position level tasks. The order of the tasks within each category was consistently varied per user to minimise any distortion as a result of familiarity with any given prototype. However, the lower level overview tasks were performed by each user first to mitigate any learning curve. Learnability and memorability will not be measured as part of this evaluation as participants were not expected to conduct tasks of significant complexity.

4 Findings

4.1 Task Completion Times

The task competition times reveal some interesting patterns relating to the efficiency of each of the prototypes (Fig. 2). It was expected that the three overview tasks that were carried out first would be easier, but there is a surprising drop in completion

times for the final three. This could be due to the participants increasing familiarity with each of the systems. What is clearer is that Prototype C is the most efficient of the three systems.

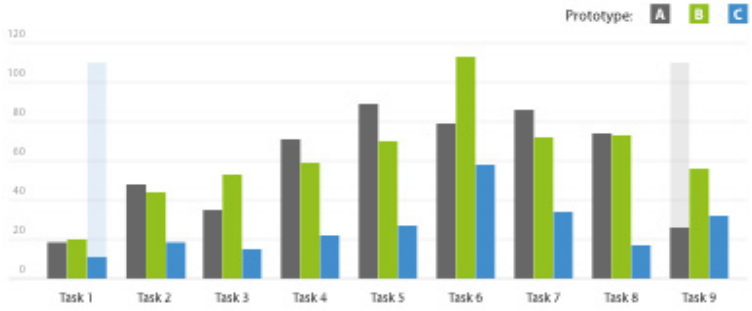


Fig. 2. A voice map measuring ease of use and satisfaction

4.2 Error Rates

There were only three errors made throughout the evaluation. While this is positive, the numbers are too small to draw concrete conclusions from. They do however raise issues that should be addressed in a next phase of design.

Two of the errors were made on Prototype C during the same task where participants were asked to select the player with the most points from the teams overview. In this view, two players (Robin van Persie and Wayne Rooney) stand out as the highest scorers because of the solid green colour of the rectangles but, as their points are so similar, it is difficult to differentiate between them. Further to this, the incorrect selection, Wayne Rooney, appears above Robin van Persie so is the first player the user will see. This lack of control of the treemap layout should be seen as a weakness and clearly demonstrates that this technique is less appropriate as a static representation of data. The error on Prototype A involved the user giving up after 45 seconds which would have been longer than any of the other tasks took to complete.

4.3 Key Observations

The observations that were made were brief in order to not disrupt the tasks that were being carried out. The observations were analysed and categorised into themes that highlight some recurring issues and behaviours throughout the evaluation.

Although all participants found the table sorting functionality of Prototypes A and B easily, three asked to be reminded of the multi-column sorting functionality. The participants that did master this functionality seemed to get there through trial and error, clicking a few times before consistently being able to perform the function correctly.

All participants seemed to find scrolling through the data in Prototypes A and B quite laborious with four confirming this in the later discussion. In this context, Tufte's principle of presenting data-rich displays is less relevant. Although the tables

have a large amount of information up front, there is a limit to how much you can display in one view, particularly on the average web browser.

The treemap supports the principle that it is easier to recognise than to recall. Using both table prototypes, users needed to remember data to make comparisons instead of being able to instantly compare relationships through size or colour. The charts in Prototype B are useful for analysing form throughout the season; they are less effective for making easy comparisons based on total player points.

4.4 Ease of Use and Satisfaction

The voice map in Figure 3 highlights some interesting insights into the participants' attitudes to the three visualisation systems. As a measure of satisfaction it seems Prototype B was significantly more successful than Prototype A. Although they compared similarly when measured by task completion times, it seems the participants reflective experience of prototype B was more positive because of the additional graphical representation of player points. This is emphasised by that fact that there was no significant difference in the perceived ease of use.

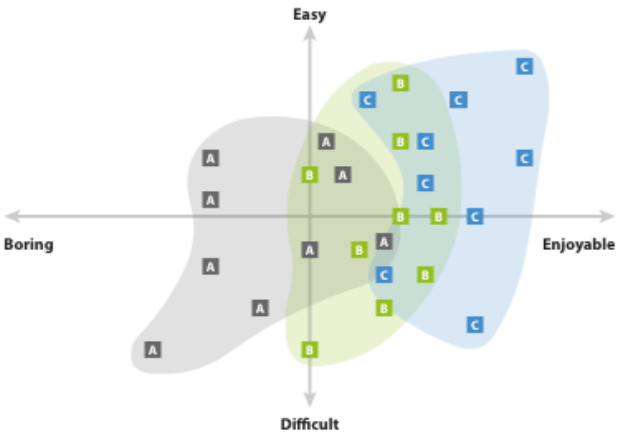


Fig. 3. A voice map measuring ease of use and satisfaction

There is also only a slight uplift in perceived ease of use for Prototype C. Although tasks carried out were completed in significantly less time than Prototypes A and B, the users still felt there was a degree of complexity to it. However, Prototype C performed better on satisfaction which could support the theory that a satisfying experiences make users more tolerant of functional complexity [13, 14].

4.5 Recommendations for Further Design

Although each prototype gave opportunities to sort or navigate through the information, there is a case for offering more effective ways of filtering unwanted information out early. Sliders are a very common interaction design component that enable users to select a value from any given scale. They can have a single control to define a

specific value or two to define a range of values. Alternatively, conversational filtering is a technique where filters are positioned in a sentence that is formed in the way that it would be spoken. This can be a more intuitive way of accessing information as it supports the user's mental model of accessing and processing information.

One of the potential challenges with the treemap is that, while it aids comparisons, size and colour are not effective ways of comparing when there is minimal variation in the data. Although the squarified treemap algorithm was created to place nodes of a similar scale together, clear differentiation isn't always guaranteed. Other techniques for organising hierarchical information that help further remove uncertainty should be considered.

One such technique is the 'Slopegraph' as created by Edward Tufte [4]. The slopegraph is a way of comparing the scale of changes over time for a list of variables. Using this technique, the variables of team performance and team spending could be ranked in ascending order and linked by lines, the angle of the lines represent the scale of differentiation. The weight of the lines would represent the scale of team spending with colour illustrating whether the differentiation is negative or positive.

This technique provides an interesting and concise alternative to the treemap layout. Where the treemap's reliance on colour and size to represent information could lead to ambiguity and confusion, the slopegraphs benefits of combining this with the angle and order of connecting lines is clear. It could be safe to assume that the two errors committed in the evaluation when comparing the player points of Robin van Persie and Wayne Rooney could have been avoided using this technique.

Another way to make interactive visualisation more intuitive is by introducing animated transitions as the user manipulates and navigates the data. Animations are a useful way of facilitating the perception of changes when transitioning between data graphics. This would be particularly relevant for the treemap which uses an interaction model of zooming and displaying information on demand.

5 Conclusions

The results of the evaluation prove, with a relative degree of certainty, that data visualisation can improve the performance and experience of fantasy football users. While this validates many of the of the established principles that informed the design, this report also proves that these principles should be used with sensitivity to the context of use. A visualisation system targeted at search and discovery as championed by Ben Shneiderman [6] will not lend itself well to high-resolution data analysis. Equally, data-rich displays of information will not perform well when deep exploration of data is required.

These findings give further weight to the theory that a positive user experience can help appease functional complexity**.** As well as reducing task times, the treemap is designed to enhance the user's emotional engagement instantly making sense of otherwise complex information by highlighting relationships and encouraging comparisons.

To take this research further, it would be useful to conduct a longitudinal study measuring the impact of data visualisation techniques over a longer period of time. This could be conducted as part of a comparative evaluation using real users

competing against each other in their natural context of use for the duration of a football season. User performance in this case would therefore be measured by competitive advantage through exploration and discovery as opposed to operational efficiency and user experience would be measured by adoption and engagement metrics.

One participant remarked that he felt empowered when using the treemap system. If fantasy football users could be genuinely empowered through the effective use of data visualisation techniques, could it then compromise competitive advantage? While fantasy football products may want to maintain an element of uncertainty in the data it makes available, this offers many possibilities for other domains.

References

1. Cleveland, McGill: Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods. Journal of the American Statistical Association 79(387) (1984)
2. Bertin: The Semiology of Graphics: Diagrams, Networks, Maps. University of Wisconsin Press, Madison (1983); Translated by Berg, W.J.: Original publisher, Mouton-Gautier. Esri Press, Redlands
3. Ware: Information Visualisation: Perception for Design, 2nd edn. Morgan Kaufmann, San Francisco (2004)
4. Tufte: The Visual Display of Quantitative Information. Graphics Press, Cheshire (1983)
5. Tufte: Envisioning Information. Graphics Press, Cheshire (1990)
6. Ahlberg, Shneiderman: Visual information seeking: tight coupling of dynamicquery filters with starfield displays. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Celebrating Interdependence, Boston, Massachusetts, United States, pp. 313–317 (1994)
7. Shneiderman: The eyes have it: A task by data type taxonomy for information visualisations. In: Proceedings of the IEEE Symposium on Visual Languages (1996)
8. Chen: Aesthetics versus functionality. Information Visualisation 1(2), 93–94 (2002)
9. Plaisant: The challenge of information visualisation evaluation. In: Proceedings of the Working Conference on Advanced Visual Interfaces, Gallipoli, Italy, May 25-28 (2004)
10. Carpendale, S.: Evaluating information visualizations. In: Kerren, A., Stasko, J.T., Fekete, J.-D., North, C. (eds.) Information Visualization. LNCS, vol. 4950, pp. 19–45. Springer, Heidelberg (2008)
11. Lam, Bertini, Isenberg, Plaisant, Carpendale: Seven Guiding Scenarios for Information Visualisation Evaluation. Technical Report (2010)
12. Bruls, Huizing, Wijk: Squarified TreeMaps. In: Proceedings of Joint Eurographics and IEEE TCVG Symp. on Visualisation. IEEE Press (2000)
13. Moere, Purchase: On the role of design in information visualisation. Information Visualisation 10(4), 356–371 (2012)
14. Norman: Emotional design - Why we love (or hate) everyday things. Basic Books, New York (2004)

A Pedestrian Navigation Method for User's Safe and Easy Wayfinding

Hiroshi Furukawa and Yutaka Nakamura

University of Tsukuba, Tsukuba, Japan
furukawa@risk.tsukuba.ac.jp

Abstract. In recent years, most of mobile phones have a function of pedestrian navigation guidance. It was reported that users sometimes feel anxiety because of low accuracy of the position estimation especially in urban area and delay of information updating. In order to reduce the anxiety, a route planning algorithm is proposed in this study, which weighs user's difficulty (or easiness) of locating own current position as well as total physical distance of courses. The difficulty is estimated by valuation functions based on the "recognizability" and "visibility" of landmarks. An experimental study conducted in real situation using a prototype system to examine and refine the model for the optimal route planning. As the result, a modified model is proposed as a promising method of route planning for user's easy wayfinding.

Keywords: navigation system, landmark, salience, cognitive model, anxiety, recognizability, visibility.

1 Introduction

However pedestrian navigation guidance is one of the widespread services provided by mobile devices [1-2], it was reported that users sometimes feel anxiety because of low accuracy of the position estimation and delay of updating of the navigation display [3]. In areas with many high buildings, the accuracy of position estimation by GPS can be low when a communication state with a satellite is harmed by the obstacles. When the precise location is not given on a navigation display, geographical information about the area is the key for users to locate their position.

The purpose of our work is to develop a pedestrian navigation aid to reduce the anxiety of users [4-5]. A route planning algorithm is proposed, which weighs user's difficulty (or easiness) of locating own current position as well as total physical distance of courses. Fig.1 depicted the basic idea of the method. We focused on the roles of useful landmarks at the locating task. Two types of valuation functions are proposed to estimate effectiveness of various landmarks. The one is a valuation function for their "recognizability" and the other "visibility."

The aim of this study is to examine and improve the models for the optimal route planning. The validity of the model was confirmed by comparing the elapsed time of user's locating task and the subjective evaluation of the anxiety by participants with

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 156–165, 2013.

the estimated results by the proposed model. Based on the results of the validity, a modified model was proposed in this study.

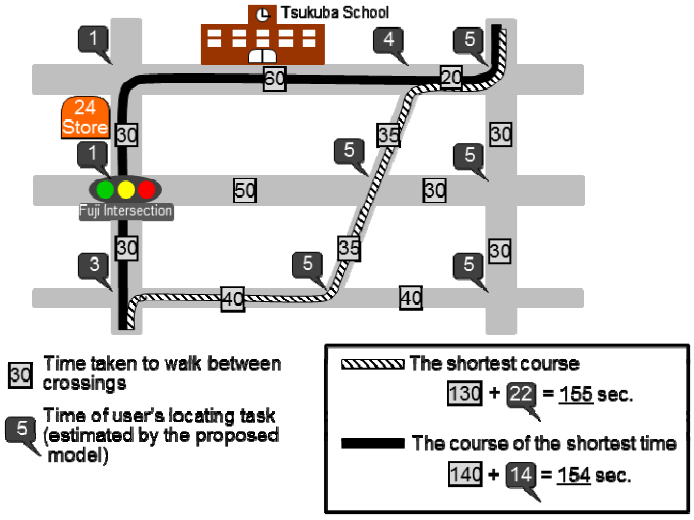


Fig. 1. The basic idea of the proposed method considering the effectiveness of landmarks, and the conventional method considering physical distance of a course

2 Related Researches

Different types of route selection methods are proposed for pedestrian navigation system to meet the needs of users. Miura and his colleagues focused on the condition about street illuminations [6]. They proposed a method evaluating the unsafe cost of the illumination and the cost of route distance. Akasaka and Onisawa proposed a method reflecting individual preference for route selection [7]. User's satisfaction degree of a road is evaluated by fuzzy measures, and used as one of the costs for route selection. In the same way, the method proposed in this paper uses multi-objective evaluation method, where the cost is defined as the easiness of locating user's current position and the distance of routes.

Some studies have conducted about evaluation methods of landmark's effectiveness for pedestrian navigation systems. Nakazawa and his colleagues proposed a method of evaluating landmark's effectiveness based on the frequency used for available navigation maps [8]. They collected guide maps from web pages and magazines, and counted frequency of the appearance of landmarks in attribute classification, such as a restaurant and a convenience store. The frequency is defined as an index of the cognitive ease of finding landmarks. An issue is that the relationships between the effectiveness of landmarks and the frequency of appearance is not clearly explained or confirmed. There is a possibility that the correlation is weak. For example, the study reported that the frequency for business complexes or apartment buildings is high [8]. On the other hand, the effectiveness of the landmarks may not be high enough to

achieve user's navigation tasks because of the low salience or low uniqueness. We are trying to develop a rational and feasible method for quantitative evaluation of the landmark effectiveness.

3 A Route Planning Method Considering Landmark's Effectiveness

Authors proposed a route planning algorithm which weighs user's easiness of locating own current position [4-5]. The easiness is quantitatively estimated using two types of valuation functions: valuation functions for the recognizability and visibility.

3.1 Recognizability of Landmarks

The recognizability is an index that shows how easy for users to find out a landmark which depicted on navigation maps. For example, it must be high when a store has a shining signboard, or when the appearance of a store is well known for users. Also, the index must be negatively correlated with distance between a landmark and user's position. The degree of influence on change of the distance must be different with different kinds of landmarks. It becomes difficult to recognize a small store like a convenience store being located in the position away from user's current location.

The valuation model for recognizability was constructed through several cognitive experiments [4]. For each type of landmarks, a regression line was determined as a valuation function for the recognizability, which shows relationships between recognizability and the distance.

3.2 Visibility of Landmarks

The recognizability index is argued about the easiness of finding a landmark assuming that it is within user's view. The visibility is defined as possibility that a landmark is in sight of users. It must be low when there are many high buildings between the landmark and the users. To consider a situation that a landmark is concealed by other buildings from user's view, two physical features are defined as factors in the visibility of landmarks: denseness of buildings around user's current position and depth of a landmark from its nearest road.

The first factor denseness is defined as a function of each base area of buildings around user's position and distance to the buildings from user's position, as described in Eq. (1).

$$denseness = \sum_{n \in N} \sqrt{A_n} / d_n \tag{1}$$

where n: a building n which is within fixed distance from user's current place, A_n: the base area of a building n [m^2], d_n: the distance from user's current place to a building n [m].

The second factor depth is defined to consider a situation that a user can see a landmark on the road which extends from his/her current location.

The valuation function for visibility is modeled using a logistic function (Eq. (2)).

$$visibility = 1/\{1 + exp[A(x - By) - C]\} \tag{2}$$

where x: the depth of a landmark from its nearest street [m], y: denseness of buildings around user's current position, A: a coefficient for the depth, B: a coefficient for the denseness, C: a bias, which represents the rapidly changing point of the function according to the depth.

To acquire data for construction of the valuation function for visibility of each type of landmarks, visual confirmation investigations of landmarks were conducted using the Street View function of Google map [4].

3.3 Valuation Function of Landmark's Effectiveness

It is thought that the easiness of locating current position at a place is equivalent to the total effectiveness of landmarks at a place in this study [4-5]. The valuation model of the total effectiveness is defined as Eq. (3), based on the valuation functions of recognizability and visibility [5].

$$CV(n) = \sum_{l=1}^{3} w_l \, C'(l) \times V(l) + \delta \tag{3}$$

$$C'(l) = 1.514 \, C(l) - 0.028 \, d_l - 2.174$$

where $CV(n)$: the total effectiveness of landmarks at a place n, l: the lth nearest landmark which is within fixed distance from a place n, $C(l)$: the recognizability of landmark l where $C(x)$ is the valuation functions of recognizability for a landmark x, $V(l)$: the visibility of landmark l where $V(x)$ is the valuation functions of visibility for a landmark x, d_l: the distance from user's current place to a building l [m], w_l: coefficients for the effectiveness of landmark l, δ: a bias. The product of the recognizability and visibility functions, $C(l)$ x $V(l)$, synthetically evaluates the easiness of finding out a landmark when it is actually within user's view. The values of the coefficients w_l are 0.958, 0.004, -0.007, and the bias δ is 0.194.

3.4 Route Planning Algorithm for User's Easy Wayfinding

In the general navigation systems, path planning is performed using the shortest path planning algorithm by Dijkstra [9]. In this study, the cost function for evaluating routes is expanded by adding elapsed time of user's locating task at each crossing to the original cost [4-5]. The elapsed time at each crossing is estimated based on the valuation function of landmark's effectiveness Eq. (3). When the cost between nodes m and n is defined as $d(m,n)$ in the Dijkstra method, the proposed cost, $d_n(m,n)$, can be described as Eq. (4).

$$d_n(m,n) = d(m,n) + dcv(n) \tag{4}$$

where $d_n(m,n)$: a cost function defined in this study, $d(m,n)$: the original cost function based on route distance, $dcv(n)$: the additional cost proposed in this study.

The $dcv(n)$ is the additional cost which considers elapsed time of user's locating task at a place n. Eq. (5) shows the relationships between the elapsed time and the total effectiveness of landmarks at a place n.

$$dcv(n) = 5.0 / \{1 + CV(n)\}, \tag{5}$$

where $CV(n)$: the total effectiveness of landmarks at a place n.

4 Cognitive Experiment

This section shows an experimental study conducted in real situation to acquire actual data for the evaluation and refinement of the proposed method described in Section 3.

4.1 Method

The main tasks of participants were achievement of navigation tasks, i.e., following an instructed route from a starting point to a destination by oneself, using a prototype pedestrian navigation system via a smartphone, and evaluation of their own anxiety during the navigation task. When they recognize that they made mistakes, they were asked to go back to the point made mistakes and restart the task.

Routes. The experiment was conducted at three different areas in Tokyo, i.e., Taitou, Setagaya, and Ikebukuro. These areas were chosen to satisfy a condition that the experimental routes include different types of circumstances. The sizes and shapes of buildings and the number of available landmarks are different from each other.

The elapsed times of navigation tasks were estimated with routes by the conventional method, considering only physical distance, and the proposed method for all possible sets of starting point and destination among each area. The five sets were selected among the three areas, where the difference in the estimated time between the conventional and proposed method conditions is large.

Prototype system. Participants were asked to walk through the assigned route with a digital map displayed on a smartphone (Sharp SO-04D). An example of the display is shown in Fig. 2. The functions implemented to the system are listed below, which are designed to simulate conditions of real navigational services (such as [1]).

- Indication of current position using information from the GPS
- Indication of a route from a starting point to destination
- Automatic periodic updating of the display
- Manual selection of the map scale
- Indication of the remaining distance

The map data were supplied by ZENRIN DataCom Co., Ltd. Several landmarks are depicted on the digital map, including ones used in the proposed route planning method, i.e., convenience stores, restaurants, bank buildings, gasoline stations, school buildings, and public buildings.

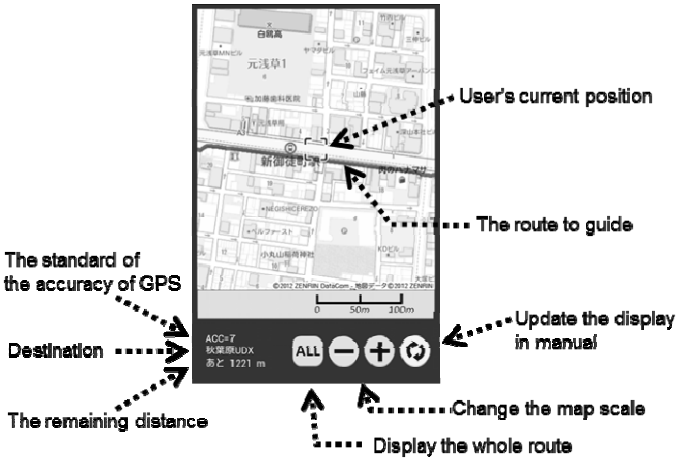


Fig. 2. An example of the display of the prototype system

Evaluation of anxiety. At selected intersections, the participants were asked to evaluate the anxiety using five-level index, where 5: "in great anxiety," 4: "in little anxiety, 3: "middle," 2: "with little assurance," and 1: "with great assurance." Because the indications were shown on the smartphone display, it was easy for the participants to achieve the evaluation task. The selection of intersections was at random manner.

4.2 Procedures

Nine paid participants, male university students, took part in this experiment. The participants were randomly divided into two groups, five in Group A and four in Group B. With each set of a starting point and destination, one group was assigned a route planned by the conventional method, and the other group assigned a route by the proposed method. In a group, the method used for planning was assigned alternately.

Total time of the experiment for a participant was between three and four hours. Total distance a participant walked was approximately 5.5 kilometers.

The data acquired in the experiments were listed below.

- the elapsed time of navigation task for a route (time for waiting for the light to change was deleted from the measured data)
- the subjective evaluation of participant's anxiety at selected intersections (using five-level index on the smartphone display)
- the reasons of the evaluation of anxiety (oral replies)

5 Evaluation and Refinement of the Model

This section shows the results of evaluation of the proposed models based on the data acquired through the cognitive experiment, and refinement of the model for the total.

5.1 Evaluation of the Proposed Model

The first evaluation is done by using the objective data from the experiment. Fig. 3 shows the relationships between the total time for user's locating task in a route estimated by the model described in Section 3, and the actual total time acquired in the experiment. The correlation coefficient between the two variables is 0.075, where the value indicated there is no dependency between them.

As the second evaluation, correlation coefficients were calculated between the elapsed time of user's locating task and the subjective data about the user's anxiety acquired in the experiment at the same intersection. The first row, the title is "The original model in Section 3", of Table 1 shows the results of every participant. The values of five participants are lower than 0.3, which can be interpreted that the dependency between the two variables is none or low.

These results show that the proposed model is not proper for estimating the valuation of the total effectiveness of landmarks or the user's anxeity.

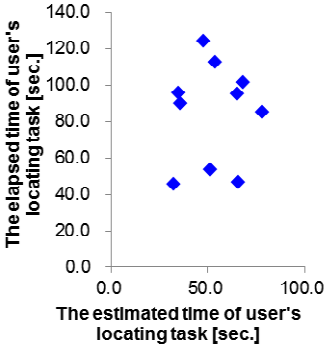


Fig. 3. The total time of user's locating task in a route estimated by the proposed model with Eqs. (4) & (5), and the actual total time as the result of the experiment

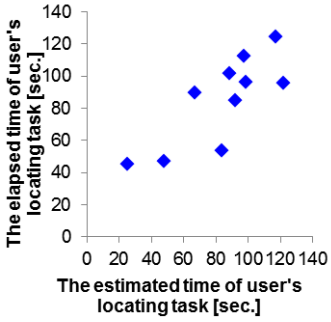


Fig. 4. The total time for user's locating task in a route estimated by the model with Eqs. (6) & (7), and the actual total time as the result of the experiment

5.2 Refinement of the Model

Two possible causes of the issue were considered in this modification process of the proposed model. The one is that the assumption about timing of user's locating task is different from actual situation. The other is possibility that the Eq. (5) is not proper model for estimating elapsed time of user's locating task at a place.

Table 1. Correlation coefficients between the elapsed time of user's locating task estimated by the proposed models and the subjective data about the anxiety at the same intersection

Group	Participants	Correlation coefficients		
		The original model in Section 3	The model with Eqs. (5) & (6)	The model with Eqs. (6) & (7)
A	a1	0.28	**0.50**	**0.46**
	a2	**0.44**	**0.59**	**0.61**
	a3	0.30	**0.37**	**0.33**
	a4	0.12	0.18	0.07
	a5	**0.36**	**0.41**	**0.40**
B	b1	**0.45**	**0.55**	**0.53**
	b2	**0.53**	**0.56**	**0.59**
	b3	0.24	**0.57**	**0.58**
	b4	0.08	**0.45**	**0.44**

The **bold number** indicates that the correlation is medium (>0.3) or strong (>0.5).

Modification Considering Timing of User's Tasks. In the interview of the participants, some reported that they did not feel any anxiety when they were walking through intersections guided to go straight. The reason is that locating tasks were not necessary at those intersections. On the other hand, the proposed route planning algorithm was developed on an assumption that users try to locate their current positions at every intersection, as described in Section 3.4.

Based on the results of the interview, a modified model described as Eq. (6) is proposed which consider the user's rule of executing a locating task at each intersection.

$$d_n(m,n) = \begin{cases} d(m,n) + dcv(n), & n \text{ Intersection } n \text{ is guided to turn} \\ d(m,n), & n \text{ Intersection } n \text{ is guied to go straight} \end{cases} \quad (6)$$

Correlation coefficients were calculated between the elapsed time of user's locating task and the subjective data about the user's anxiety at the same intersection acquired in the experiment. The second row ("The model with Eqs. (5)&(6)") of Table 1 shows the results for every participants. The values of seven participants are higher than 0.4, which suggests that the correlation between the two variables is medium or strong.

An independent two-sample *t*-test was conducted to compare correlation coefficients in conditions with the original model and the modified model with Eq. (6). There was a significant difference in the correlation coefficients in the previous model (M = 0.31, SD = 0.15) and the modified model (M = 0.46, SD = 0.13) conditions; $t(16) = 2.12$, $p = 0.035$ (<0.05).

These results suggest that the modified model with Eq. (6) is more proper than the previous model to estimate the elapsed time of user's locating task and the anxiety.

Modification on the Model for Elapsed Time of User's Task. The aim of study in this section is to modify Eq. (5). The basic assumption of Eq. (5) is that the relationships between the elapsed time and the effectiveness is monotone decreasing, i.e., as the effectiveness is getting higher, the time getting shorter.

Several monotone decreasing functions were evaluated as the alternatives using correlation coefficients with the measured data. The results of evaluation suggest that Eq. (7) is a proper model for the estimation.

$$dcv(n) = \begin{cases} 23.0, & CV(n) < 1.0 \\ 23.0/CV(n), & otherwise \end{cases} \tag{7}$$

Fig. 4 shows the relationships between the total times for user's locating task in a route estimated by the model (Eqs. (6) and (7)), and the actual total time measured in the experiments. The correlation coefficient between the two variables is 0.83, where the value indicated the correlation is very strong. On the other hand, the correlation coefficient with the Eqs. (5) and (6) is 0.68, which is also strong but less. These results suggest that the modified model using Eqs. (6) and (7) is proper to estimate the elapsed time of user's locating task.

Correlation coefficients were calculated between the elapsed time of user's locating task and the subjective data about the user's anxiety at the same intersection acquired in the experiment. The third row ("The model with Eqs. (6)&(7)") of Table 1 shows the results for every participant. An independent two-sample *t*-test was conducted to compare correlation coefficients in conditions with the modified model with Eqs. (5) & (6) and Eqs. (6) & (7). There was no significant difference in the correlation coefficients in the Eqs. (5) & (6) model (M = 0.46, SD = 0.13) and the Eqs. (6) & (7) model (M = 0.45, SD = 0.17) conditions; $t(15) = 2.13$, $p = 0.80$.

Summary. Even though the superiority in estimating user's anxiety is not confirmed, these results suggest that the modified model with Eqs. (6) & (7) is a promising method of route planning in this study.

An issue should be concerned is that an individual difference in the estimation. The low correlation coefficients of a participant (a4 in Table 1) indicate that there are some other meaningful factors not concerned in the proposed model. Nevertheless, it is expected that the proposed method is proper for supporting navigational tasks of most of users.

6 Conclusions

In order to reduce anxiety of a pedestrian navigation user, a route planning algorithm is proposed in this study, which weighs user's difficulty (or easiness) of locating own current position as well as total physical distance of courses.

Recognizability and *visibility* of landmarks are defined as basic factors which specify the effectiveness of landmarks in user's position estimation. The cost function for evaluating routes is expanded by adding elapsed time of locating intersections to the original cost of route distance. The elapsed time at an intersection is estimated based on the valuation functions of effectiveness of landmarks. When a user inputs his/her destination to navigation system, a route with the lowest cost is selected to support user to reach the goal without anxiety.

The aim of this study is to examine and refine the models for the optimal route planning. The validity of the valuation functions was confirmed by comparing the elapsed time of user's locating task and the subjective evaluation of the anxiety with the estimated results by the proposed model. As the result, this study proposes the modified model as a promising method of route planning for user's easy wayfinding. It is expected that the proposed method is proper for supporting navigational tasks of most of users.

Future tasks of this study are following:

1. The first task is a study on the individual difference in the relationships between the elapsed time and use's anxiety of locating tasks must be our next theme to improve the usefulness of the route navigation method.
2. The second task is the validation test of the pedestrian navigation guidance with participants of different conditions. The aim of the test is to confirm that the guidance can reduce anxiety of users actually in real situations.

Acknowledgment. This work has been partially supported by Grants-Aid for Science Research 23500109 of the Japanese Ministry of Education, Science, Sports and Culture.

References

1. Navitime Japan (in English), https://www.navitime.co.jp/en
2. EZ Navi Walk (in Japanese), http://www.au.kddi.com/ez_naviwalk/
3. Tatenami, T., Matsushita, D., Munemoto, J.: Transition of level of anxiety in wayfinding with pedestrian navigation system on mobile telephone. Journal of Architecture and Planning 608, 59–64 (2006) (in Japanese)
4. Furukawa, H., Uto, H.: A quantitative evaluation method of landmark effectiveness for pedestrian navigation. In: UKSim 14th International Conference on Computer Modelling and Simulation. IEEE Computer Society (2012)
5. Furukawa, H., Nakamura, Y.: Pedestrian navigation system based on quantitative estimation of landmark's effectiveness. In: The 4th International Conference on Applied Human Factors and Ergonomics 2012, pp. 391–400. USA Publishing (2012)
6. Miura, H., Takeshima, S., Matsuda, N., Taki, H.: A study on navigation system for pedestrians based on street illuminations. In: König, A., Dengel, A., Hinkelmann, K., Kise, K., Howlett, R.J., Jain, L.C. (eds.) KES 2011, Part III. LNCS, vol. 6883, pp. 49–55. Springer, Heidelberg (2011)
7. Akasaka, Y., Onisawa, T.: Personalized pedestrian navigation system with subjective preference based route selection. In: Ruan, D., Hardeman, F., Meer, K. (eds.) Intelligent Decision and Policy Making Support Systems. SCI, vol. 117, pp. 73–91. Springer, Heidelberg (2008)
8. Nakazawa, K., Kita, N., Takagi, K., et al.: A Dynamic Map Based on Landmark's Visibility. Transactions of Information Processing Society of Japan 49, 233–241 (2008) (in Japanese)
9. Dijkstra, E.W.: A note on two problems in connexion with graphs. Numerische Mathematik 1, 269–271 (1959)

The Effect of Information Quantity on Cbp Interface in the Advanced Nuclear Power Plant

Min-Chih Hsieh and Sheue-Ling Hwang

National Tsing Hua University, Hsinchu, Taiwan 300
g9674019@cycu.org.tw

Abstract. Computer-based procedures (CBP) are widely used in advanced nuclear power plant (NPP). In order to improve operators' performance and decrease the probability of human errors, several different CBP interface layouts have been provided by previous studies. However, there has been little research discussing the information quantity that present on the CBP interface. This study is to examine operator performance in using computerized procedures from an information quantity point of view. A simulated CBP system was developed to present three different information present styles, each with different task complexity. The results indicated that medium information quantity of CBP interface was better than other level of information quantity.

Keywords: Computer-based procedures; Nuclear power plant; interface; information quantity.

1 Introduction

The procedures of system control that guide operators in performing tasks are widely used in nuclear power plant (NPP) domain (Husseiny et al., 1989; Xu et al., 2008). Computerized procedures (CP) systems in a nuclear power plant (NPP) provide instructions to guide operators in monitoring, decision-making, and controlling the plant. The CP systems have been accepted by various regulatory authorities and are in use at several NPP around the world (Yang et al., 2012). The results of other research studies of CP systems show that CP systems can provide some performance benefits, such as tasks can be performed more quickly; mental workload can be reduced and minimized; fewer errors may be made in transitioning through or between procedures (NUREG/CR-6634; NUREG/CR-6749; Portmann and Lipner, 2002; O;Hara et al., 2003; Yang et al., 2012).

Computer-based procedure (CBP) was a human-system interaction technology that was found to be potentially safety significant (Yang et al., 2012). In order to improve operators' performance and decrease the probability of human errors, several different CBP interface layout have been provided by previous studies (Jung et al., 2004; Xu et al., 2008; Carvalho et al., 2008; Hong et al., 2009).

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 166–173, 2013.

Regarding to the control interface of CBP studies, Wourms & Rankin (1994) pointed out that flowchart is better than text in the procedure format because of its ability to relate procedure elements explicitly. Jun et al. (2004) published coherently coupled frames that reduced the number of navigation, and flowchart and success tree were used where total of the process information was shown on the interface. Xu et al. (2008) compared two mainstreams graphical computerized procedures such as flowchart procedures and success tree, the results showed that flowchart procedure was better than success tree. Because the flowchart procedure resulted in a lower error rate without introducing a longer task time and more workload. Carvalho et al. (2011) developed a new CBP interface, and compared it with existing interfaces. The results showed that the time spent by the operator to identify two accidents, through the new interface, was faster than the existing interface. Hong et al. (2009) also developed a computerized procedures system to assist operator executing the procedure tasks.

No matter what format that the procedure information presenting on CBP, however, there has been little research discussing the information quantity that present on the CBP interface. Operators who work in the NPP main control room might be influenced by the information quantity from the computer screen (Wickens, 2000). The objective of this study, therefore, is to examine operator performance in using computerized procedures from an information quantity point of view.

This paper is organized as follows: Section 2 describes the experiment design; Section 3 presents the results of this experiment; Section 4 presents discussions and conclusions drawn from the study.

2 Methodology

The information quantity issues of CBP interface were investigated by experiment with subject participation. The details about independent and dependent variables, the experiment process, and the information quantity of CBP interface will be introduced in this section.

2.1 Independent Variables

Two independent factors were considered in this experiment: the information quantity of CBP interface and task complexity. Information quantity of CBP interface consisted of three levels and task complexity consisted two levels, as listed in Table 1. A between- and within-subjects mixed experiment design was adopted where a subject would participate in the experimental condition of one presentation style and both task complexity levels.

Table 1. Factors and levels in the experiments

Factor	Level 1	Level 2	Level 3
Information quantity	High	Medium	Low
Task complexity	Complex	Simple	-

2.2 Information Quantity of CBP Interface

Three information quantity of CBP interface (high, medium, and low) were adopted in this experiment. The measurement in information quantity used in this experiment lay in the number of events on the CBP interface, which can be expressed by the formula as follows (Wickens, 2000):

$$H_s = \log_2 N. \quad (1)$$

The information conveyed by an event H_s, in bits, and N is the number of alternative. For example, high information quantity in this experiment with at least eight events' information was shown on the CBP interface in each step, $\log_2 8 = 3$ bits. Medium information quantity in this experiment with at most four events' information was shown in each step, $\log_2 4 = 2$. Low information quantity in this experiment only have one event information was shown on the CBP interface, $\log_2 1 = 0$. Each level of factors included flowcharts and brief instructions. The flowcharts located on the left side of the computer screen displayed all steps of the procedure with the current step highlighted. The flowcharts located in the center of the computer screen displayed the detail structure of the current step. System states were located in the right side of the display.

2.3 Task Complexity

In this study, the degree in task complexity lay in the numbers of steps and parameters in the procedures (Xu et al., 2008). The procedure with high complexity had more steps and parameters than the one with low complexity (table 2).

Table 2. Procedures with complex and simple task

Task complexity	Number of steps	Number of parameters
Complex	9	99
Simple	5	45

2.4 Dependent Variables

Both objective and subjective indexes were used to evaluate operation performance. The objective indexes include operation time and secondary task performance; the National Aeronautics and Space Administration task load index (NASA TLX) was used to measure the subjective workload.

2.5 Participants and Experimental Task

Twenty-one college students from National Tsing Hua University were recruited as participants in this experiment. The average age of the participants was 25.5 years and ranged from 22 to 29 years. All participants were randomly assigned into one of three different information presentation styles. Half of the participants in each presentation

style start the formal experiment from complex task, others participants start from simple task.

The CBP interface was simulated using a computer program developed by Adobe FLASH CS6, as shown in Fig. 1.

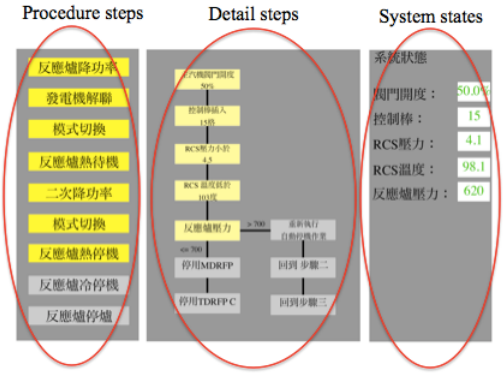


Fig. 1. Interface of the experiment platform (Information quantity: medium; Task: complex)

2.6 Experimental Procedure

The participants were told to try their best to perform the task and avoid any mistakes during the formal experiment. During the primary task, participants are required to detect signal as a searching task. This signals which are randomly showed on the board in front of the participant and the response time of participants were recorded. If the participants miss it and do not detect the signal in 15s, it is recorded as a miss.

Before the formal experiment, the participants needed to listen to the experiment explanation from the experimenter and practiced the CBP system 5 to 10 minutes. At the start of the formal experiment, the participants' operation time was recorded. After the participants finished an experiment treatment, they were asked to take the first NASA-TLX questionnaire and took a break and then continued to execute the other experiment treatment. Finally, the participants took the second NASA-TLX questionnaire and finished the experiment.

3 Results

In this experiment, all statistical analyses were carried out with MINITAB 16. Multivariate analysis of variance (MANOVA) was computed for the operation time, secondary task performance, and workload.

The results of MANOVA indicated that information quantity (Willks' $\lambda = 3.743$, $P < 0.01$), and task complexity (Willks' $\lambda = 10.289$, $P < 0.01$) had significant effects on the dependent measures, as shown in Table 3 and Table 4. There was no significant interaction between information quantity and task complexity on operation time, secondary task performance, nor workload (Willks' $\lambda = 3.484$, $P < 0.05$).

Table 3. MANOVA for information quantity

Criterion	Test Statistic	F	P
Wilks'	0.50943	3.743	0.003
Lawley-Hotelling	0.87084	3.919	0.003
Pillai's	0.53751	3.553	0.005
Roy's	0.74759		

Table 4. MANOVA for task complexity

Criterion	Test Statistic	F	P
Wilks'	0.47564	10.289	0.000
Lawley-Hotelling	1.10243	10.289	0.000
Pillai's	0.52436	10.289	0.000
Roy's	1.10243		

With regard to the individual dependent variables, information quantity significantly influenced the operation time ($F = 9.66$, $P < 0.01$), secondary task performance ($F = 5.37$, $P < 0.05$), and workload ($F = 5.41$, $P < 0.05$).

Figure 2 shows the operation time under each information quantity. Comparing the operation time between these three information quantities demonstrated that the operation time in the high and medium quantity was faster than that in the low quantity.

The secondary task performance under each information quantity is showed in Figure 3. The secondary task performance in the medium quantity was significantly better than that in the high and low information quantity.

Figure 4 shows the workload under each level of information quantity. Comparing the workload between these three information quantities demonstrated that the workload in the medium quantity were lower than that in the high and low quantity.

With regard to the individual dependent variables, task complexity significantly influenced the operation time ($F = 32.47$, $P < 0.01$), but not the workload nor secondary task performance.

Comparing the operation time between these two tasks complexity demonstrated that the operation time in the simple task was faster than that in the complex task.

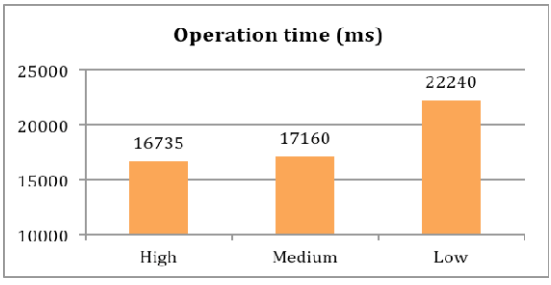


Fig. 2. Main effect plot for operation time

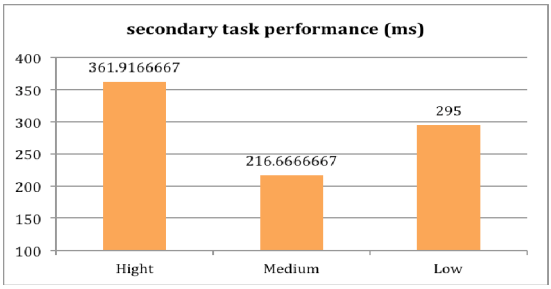


Fig. 3. Main effect plot for secondary task performance

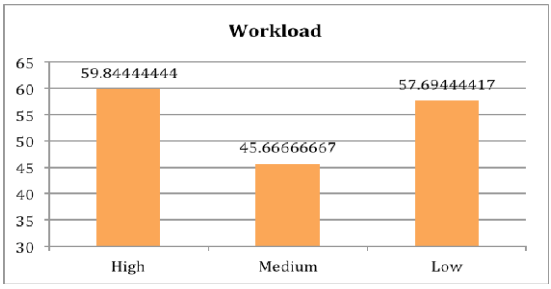


Fig. 4. Main effect plot for workload

4 Discussion and Conclusion

In this study, an experiment was designed to explore how the information quantity of CBP interface and task complexity influence the performance of operators from the criteria of operation time, secondary task performance, and subjective workload.

The information quantity significantly influenced the operation time, secondary task performance, and workload. When the CBP interface with high information quantity, the participants need to spend more time and pay more attention on the task, and may ignore the secondary task and increase mental workload. On the other hand, when the CBP interface with low information quantity, the participants need more actions to click on procedure steps producing a detailed system state of the procedure step, and thus the participants spent more time on the low information quantity task. The task complexity significantly influenced the operation time, in this experiment, and complex task corresponded to longer task time.

Three levels of information quantity were compared in this experiment. Medium information quantity could cause a significantly lower workload than high and low information quantity. Furthermore, Medium information quantity resulted in the highest secondary task performance. There are two reasons to explain the result of this experiment. First reason is that it decreases the information quantity of detailed system state from high information quantity of CBP interface. The other reason is that

medium information quantity of CBP interface decreases the number of click CBP interface from low information quantity of CBP interface. Therefore, the participant could identify system state more efficiently, and induce lower mental workload. These results suggest that medium information quantity of CBP interface is better than other level of information quantity.

The interface in an advanced NPP main control room are modernized fully integrated digital design, where CBP have a range of capabilities that may support operators in controlling the NPP (Yang et al., 2012). This study verified the information quantity of CBP interface, including operation time, secondary task performance, and mental workload. In conclusions, the results of this study may be applied on the advanced NPP, as well as be generalized to the design of information display of other computerized system, such as process control system, air traffic control system and flight management system.

Acknowledgments. Partial of this study was supported by National Science Council, Taiwan, under project No. NSC101-2221-E-007-013-MY2.

References

1. Husseiny, A.A., Sabri, Z.A., Packer, D., Holmes, J.W., Rodriguez, R.J.: Operating procedure automation to enhance safety of nuclear power plants. Nucl. Eng. Des. 110(3), 277–297 (1989)
2. Xu, S., Song, F., Li, Z., Zhao, Q., Lou, W., He, X., Salvendy, G.: An ergonomics study of computerized emergency operating procedures: Presentation style, task complexity, and training level. Reliability Engineering and System Safety 93, 1500–1511 (2008)
3. Yang, C.W., Yang, L.C., Cheng, T.C., Jou, Y.T., Chiou, S.W.: Assessing mental workload and situation awareness in the evaluation of computerized procedures in the main control room. Nuclear Engineering and Design 205, 713–719 (2012)
4. NUREG/CR-6634, O'Hara, J., Higgins, J., Stubler, W., Kramer, J.: Computer-based Procedure Systems: Technical Basis and Human Factors Review Guidance (NUREG/CR-6634). U.S. Nuclear Regulatory Commission, Washington, DC. (2000)
5. NUREG/CR-6749, O'Hara, J.: Integrating digital and conventional human system interface technology: Lessons learned from a control room modernization program (NUREG/CR-6749). U.S. Nuclear Regulatory Commission, Washington, DC. (2002)
6. Portmann, F., Lipner, M.H.: An Operational Model for Using a Computerized Emergency Operating Procedures System. Modern Power Systems (2002)
7. O'Hara, J., Pirus, D., Nilsen, S., Biso, R., Hulsund, J.E., Zhang, W.: Computerisation of procedures lessons learned and future perspectives. OECD HALDEN REACTOR PROJECT. HPR-355 (2003)
8. Jung, Y., Seong, P.H., Kim, M.C.: A model for computerized procedures based on flowcharts and success logic trees. Reliability Engineering and System Safety 26, 351–362 (2004)
9. Carvalho, P.V.R., dos Santos, I.L., Gomes, J.O., Borges, M.R.S., Guerlain, S.: Human factors approach for evaluation and redesign of human-system interfaces of a nuclear power plant simulator 29, 273–284 (2008)

10. Hong, J.H., Lee, M.S., Hwang, D.H.: Computerized procedure system for the 1400 simulator. Nuclear Engineering and Design 239, 3092–3104 (2009)
11. Wourms, D., Rankin, W.: Computer based procedure (Tech Report CSERIADC-RA-94-002). Wright-Patterson Air Force Base. Crew Systems Ergonomics Information Analysis Center, OH (1994)
12. Wickens, C.D., Hollands, J.G.: Engineering psychology and human performance. Prentice-Hall Inc., NJ (2000)

Development of Communication Support System at Mealtimes Using Tabletop Interface

Junko Itou, Shizuka Takao, and Jun Munemori

Faculty of systems Engineering, Wakayama University, 930, Sakaedani,
Wakayama 640-8510, Japan
{itou,munemori}@sys.wakayama-u.ac.jp

Abstract. In this research, we aim to develop a system to support communication during meal times. It is reported that the communication skills of young people have declined because of decrease of opportunity to speak face to face. We focus on the meal that is essential for everyday life as a place for face-to-face communication. The proposed system uses tabletop interface as a dining table to provide intuitive control without additional burden. Users can decorate the table and dishes with several digital items by only touching the surface of the table. Furthermore the system displays various visual events randomly on the table according to the condition of the items. By using visual effects, we realize the support to provide improving communication and enhancing the enjoyment of the meal. From the result of experiment, it was cleared that users could take a cue of conversation through use of the system.

1 Introduction

We propose a system that people can enjoy communication during their meal using table top interface. The research for approximately 400 college and high school students by the Ministry of Health, Labor and Welfare[1] reported the decline in communication skills of Japanese young people. One-third of the target of the research felt that the ability to communicate was insufficient. As the main cause, the report pointed out the spread of mobile phones and the advancement of information networks.

In face-to-face communication such as daily conversations, people convey linguistic information with non-verbal information including gestures and facial expressions. On the other hand, the only verbal information is exchanged in communication via networking tools. It is not necessary to meet other people and it will reduce the opportunity to express and to construe the nonverbal expressions. This situation is one of factor of decline in communication skills. Therefore the support of face-to-face communication is a very important issue. In order to support making an opportunity of face-to-face communication and activating the communication, we focused on the meal that everyone always do in daily life.

We adopt a table top interface as a dining table to support communication during meal times. The table can provide intuitive control so users can use the system easily

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 174–182, 2013.

while eating. The existing researches[2][3] provide the fun of mealtime communication by sharing photographs or decoration items using cameras and a projector. However, it is limited to operate the table because the user's head or his/her hand blocks the camera's view and only one person controls the table at a time. We realize an interactive table which many users can control freely and easily at a time so that the users enjoy their meal times by viewing visual effects. E-Table[4] is an interactive ordering system that diners can order food and drinks directly from a menu projected on to their table in their hands in a restaurant and bar. Diners also can search and browse information of the dishes or the local neighborhood of the restaurant until a wait staff serves, and call a taxi during their meal times using this table. Whereas E-table focuses on rationalization in business and fun to pass the time until dishes are served, our system support to make opportunity to speak face to face among diners by displaying visual effects including accidental visual events on the table.

This paper is organized as follows: in section 2, we will describe the related services on support of mealtime communication and interactive systems for dinners using table top interface. In section 3, we explain our system that supports to initiate the conversation starter and enhancing the enjoyment of meal time communication. A validation test of our system will be given in section 4. Finally, we discuss conclusions and a future work in section 5.

2 Related Research on a Support System of Meal Time Communication

There have been a number of communication support systems at mealtimes using a camera, a projector, a meal plate and a table. We discuss an effective approach to support meal time communication aiming to make an opportunity to speak face to face in this section.

pHotOluck[2] is a mealtime communication support system that consists of dishes, a projector and a camera. The projector mounted to the ceiling displays photographs that users have taken previously, on a meal plate on a dining table. The camera recognizes three markers on the plate and estimates the position to project an image even if users change the plate's position and its direction. However, this system has a problem with operability because users cannot act in such a way as to cover the markers on the plate form the camera by their hands or heads. Furthermore, people hold chopsticks, a knife, a fork and a spoon in their hands during meal time so it is necessary to achieve more intuitive interface that users don't have to pay mind to operate.

Mori et al. has proposed an augmented reality system in order to make meal time fun[3]. This system takes a movie of a dish on a dining table with a video camera. The system analyzes its colors and calculates colors to make the dishes look more delicious based on the analysis. The result is projected onto the dish as a decoration of its rim. In addition, the image that a cock designs with illustrations and messages to decorate the dining table, the meal plate and the dish in the kitchen is also projected. The cock can change the contents which the projector shows according to remaining

amount of the dish. This research concluded that the visual effects improved palatability and pleasure during meal time. On the other hand, the eater itself cannot change the rim decorations and the illustrations. Furthermore it is limited that the meal plate the image is projected on is only one. Therefore this system is unsuitable for the situation that several persons sit around the dinner table and communicate face to face.

Compurants Ltd. has released an interactive ordering system 'E-table' with a table top interface[4][5]. Dinners can order foods and drinks directly touching the surface of the table where an illustrated food and drinks menu is displayed on. Dinners also can set their table ambience and watch information on the ordered dishes or the local neighborhood of the restaurant until the ordered dishes are served. Anyone can operate the system intuitively by adopting table top interface. As just described, E-table helps dinners and the restaurant staffs in the point of rationalization in business and service. On the other hands, this system doesn't play a role strongly in supporting face-to-face communication rather E-table decreases the opportunity for conversation among diners and staffs because the dinners can enjoy the meal time on his/her own.

Based on the researches and services on support of meal time communication, we aim to develop an interactive system that makes users enjoy a meal time and communication. For support face-to-face communication during meal time, we employ table top interface and several visual effects so that several users can operate at the same time and the system trigger a conversation between them.

3 System

3.1 Goal

Our goal is development of communication support system that users can operate easily and intuitively while they eat. The system also provides opportunities of conversations and makes users enjoy the mealtimes. In order to achieve this goal, we design the system as follows. First we adopt a table top interface as a meal table to reduce user's burden in term of use during a meal. Secondly, the system projects various kinds of visual effects on the table including in digital decorations with motions, a background image and a topic presenting character. Users can move freely the digital decoration items within the table even if on the dishes. Finally, several kinds of special events occur randomly based on the state of the digital decoration items displayed on the table so that the event livens up the table and it gives an opportunity of conversation among users.

3.2 System Structure

In this system, we use Diamond Touch Table (hereinafter referred to as 'DTT') as a table top interface[6]. DTT can identify up to four persons at a time. In addition, cameras and markers are not required for recognition of users' action on the table. Fig.1 shows the overview of our system. An image from the projector is displayed on the screen that is the surface of DTT. Users treat DTT as a dining table and sit around

DTT as shown Fig.2 and Fig.3. Employing DTT which provides intuitive controls reduces users' burdens and makes it possible to project many types of visual effects on the dishes and the table.

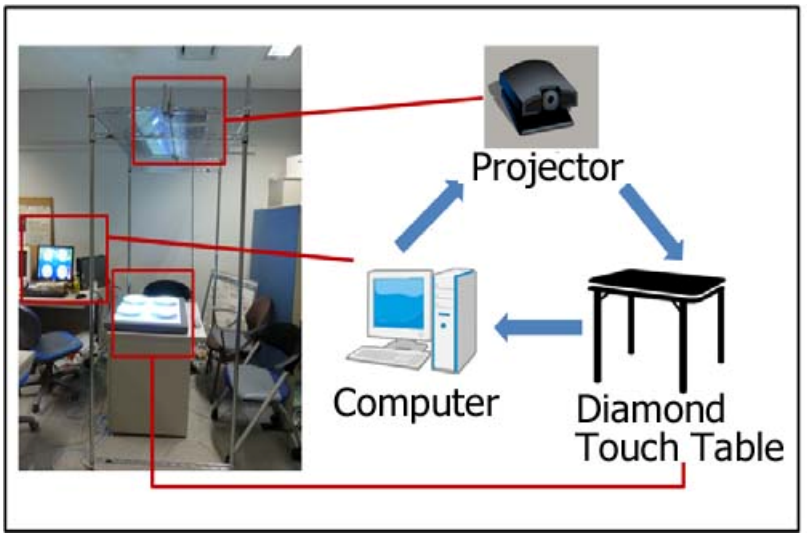


Fig. 1. Overview of proposed system

Fig. 2. Dishes and DTT as a meal table

Fig. 3. The table and users during meal times

3.3 Contents of Visual Effects

Fig.4 illustrates the image on the table. Users can manipulate the DTT by touching the surface of this table. The four white circles in Fig.4 mean the positions of meal plates. This system is considered to be used by up to four persons at same time so the buttons for operations are placed in each four corner of the table. Users can press the buttons only in front of themselves. This system provides several types of operations.

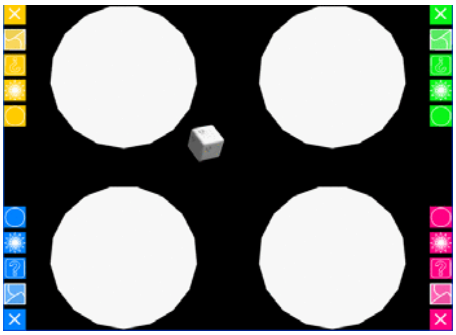


Fig. 4. An example of a projection image

Fig. 5. Decorations with flower items

Fig. 6. Change of a background image and projection of an image on a dish directly

Decoration Function

Users can decorate the table and dishes with flowers and circles. After touching the corresponding button and any position on the table, then a flower item is displayed. Any users can move the items freely by drag on the table. The color of these items is fixed depending on the seat. When a user wants to use the other color item, the user asks for the other users to produce the item and to move it. An example of a decorated dish with various color flowers and circles are shown Fig.5.

Background Image

Users can select the background image of the table from black to 10 different patterns as if they change the table cloth. We aim that the users will communicate when they want to change the background, and that the background image enhances the mood of their meal. The sunset image is used as a background image in the left image of Fig.6. As shown in the right image of Fig.6, the image is directly projected on the dish. Users can switch the display if the background image has an overlap with the dish.

These operations are performed by only touching the surface of the table so users can control even if with non-dominant hand.

Topic Presenting Character

The white cube object displayed in the center of Fig.4 is a character that provides topics of conversation. The character shows today's news and trivia in its balloon and it has 25 patterns of the message. The character constantly moves and rotates within the table as all users can read the message.

Visual Event Based on Items

Various events may occur at random according to the condition such as the item number of each color circles and flowers that are displayed on the screen. The visual event has three types. The first one is the event that special items are displayed. An example of this event is shown in Fig.7. Four big flowers are blossoming at the center of dishes in the left lower image of Fig.7.Users can move these items as with the other normal items.

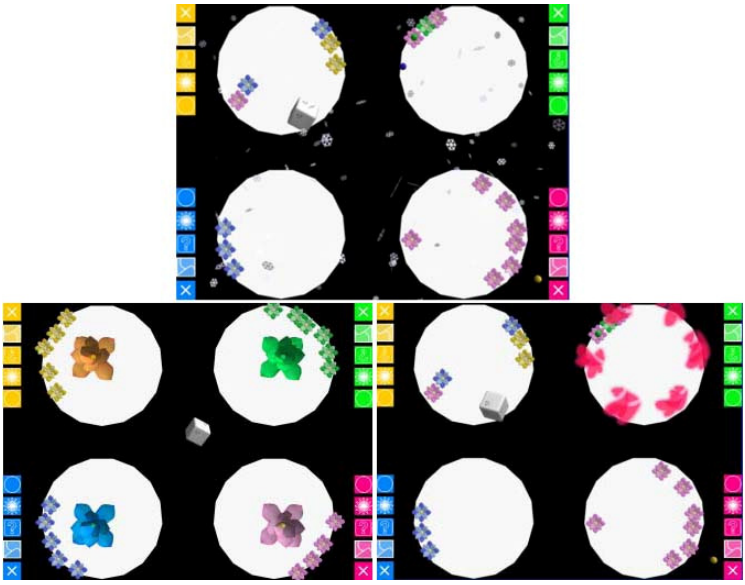


Fig. 7. Examples of visual events

Fig. 8. Projections of visual events on a dish and the table

The second type of the visual event is that some decorations with motions are displayed around the rim of the dishes. There are 27 kinds of patterns in this event. The right lower image of the Fig.7 is an example of this visual event. The result of the projection on a dish is shown in Fig.8. Some big hearts are projected in blinking and rotating around the rim of the dish.

The third type of the visual event is shower of blossoms or snow crystals. This effect covers the whole table top and it has four patterns. The top image of Fig.7 shows the shower of snow crystals on the table. The example image of projection of this event is the right image in Fig.8. The all sizes snow crystals are displayed on the table and the dishes.

By using these visual effects and decorations, we realize the support to provide opportunities of communication and enhancing the enjoyment of the meal.

4 Experimental Result

We performed an evaluation experiment using proposed system to investigate how our system effects on communication during meal times. The experimental subjects were 24 persons in their twenties divided into six groups. The three groups are consisted of people who had met for the first time. We instructed each subjects to sit at the table and to eat spaghetti and fruits. The meal time is 30 minutes. The subjects could talk freely during their meal. The conversation topic was not limited. We recorded the scene of the meal by a camera.

The result of a questionnaire on uses of the system is described in Table 1. The numbers of the Table 1 means the number of subjects who answered the each degree. The number 1 means strongly disagree and the number 5 means strongly agree. By comparing the result of item (i) and item (ii), we obtain higher value in proposed system than the usual meal. According to (iii), it is cleared that most subjects more enjoyed their meal using our system than usual meal even contains first meeting persons. The five subjects who answered “I don’t have any idea” or “It is troublesome” for the question “What do you think for your daily meal? Do you enjoy it?” selected the number 5 on the questionnaire item (i).

Table 1. Questionnaire result on use of the system

Questionnaire item	1	2	3	4	5
(i) You enjoyed your meal using the system.	0	0	0	7	17
(ii) You enjoy your usual meal with your friends.	0	0	1	10	13
(iii) This meal was more enjoyable than usual meal	0	0	0	12	11

Table 2. Questionnaire result on the support of communication

Questionnaire item	1	2	3
(iv) This system provided conversation material.	1	9	14
(v) You enjoyed this meal with first meeting persons.	1	6	5

Table 3. Analysis result of video data

		With friends			With strangers		
	Group	A	B	C	D	E	F
During eating	Time (minute)	16	10	28	28	23	18
	Time without speaking (second)	47	8	5	59	20	13
	Number of conversation triggered by the system	4	8	18	22	23	16
	Number of operation of the table	73	48	121	252	128	63
After eating	Time (minute)	14	20	2	2	7	12
	Time without speaking (second)	32	0	0	0	0	5
	Number of conversation triggered by the system	6	23	0	0	10	12
	Number of operation of the table	39	174	8	18	81	77

Table 2 is the result of the questionnaire on the communication support. The number 1 means 'not at all', the number 2 corresponds to 'sometimes' and the number 3 means 'many times'. As shown in Table 2, 23 subjects answered that they could get the conversation material more or less. In addition, 92% of the subjects whose group consisted of first meetings answered that the meal became enjoyable by using this system as shown in Table 2 (v).

Table 3 shows the analysis result of video data. The group A, B and C consists of friends and the group D, E and F consists of first meeting persons. The 86% of subjects who answered they felt gap in a conversation during the mealtimes could resolve the situation by using this system. Totally the time without speaking after eating is shortens the time during eating in all groups. On the other hand, the first meeting groups tended to operate the table more frequently after eating. Group C took time to eat their meals as much as Group D did, however the time without speaking and the number of operating the table totally differed. By comparing group C and group D, the members of the first meeting groups often touched the table and tended to devote themselves actions meanwhile declining speaking time. They also answered that this system made the opportunity to speak by providing the topic materials such as exchanging flower items and displaying random events. So we can confirm this system support the conversation among the subjects.

An important thing to note, by analysis of video images taken in several groups, it revealed that users often moved flower items and decorated dishes of their partners with conversations. Especially, the person who finished his/her meal decorated the plate of a user who still ate. They had conversation about the action and the result of the decoration.

5 Conclusions

In this article, we proposed a support system for communication at mealtimes aiming to support the activation of face-to-face conversation. We applied DTT as a table for meals and designed the system to provide an enjoyment of meals. From the experiment, it was cleared that users could take a cue of conversation through user of the system and responses from the system even contains first meeting.

References

1. The Ministry of Health, Labor and Welfare : Youth Employability Support Program (2009) (Japanese site), http://www.mhlw.go.jp/bunya/nouryoku/yes/
2. Nishimoto, K., Amano, K., Usuki, M.: pHotOluck: A Homeuse Table-ware to Vitalize Mealtime Communications by Projecting Photos onto Dishes. In: Horizontal Interactive Human-Computer Systems (TABLETOP 2006), pp. 9–16 (2006)
3. Mori, M., Kurihara, K., Tsukada, K., Siio, I.: Dining Presenter: Augmented Reality system for a dining tabletop. In: Supplemental Proceedings of the 11th Ubicomp 2009, pp. 168–169 (2009)
4. Compurants Ltd. : "E-Table", http://www.e-table-interactive.com/index.html (accessed February 19, 2013)
5. Compurants Ltd. : inamo restaurant, Soho, London - interactive oriental fusion restaurant and bar, http://www.inamo-restaurant.com/ (accessed February 19, 2013)
6. Dietz, P., Leigh, D.: DiamondTouch:A Multi-User Touch Technology. In: Proceedings of UIST 2001, pp. 219–226 (2001)

Improving User Performance in Conditional Probability Problems with Computer-Generated Diagrams

Vince Kellen, Susy Chan, and Xiaowen Fang

DePaul University, College of Computing and Digital Media, Chicago, IL, USA
vkellen@uky.edu, {schan,xfang}@depaul.edu

Abstract. Many disciplines in everyday life depend on improved performance in probability problems. Most adults struggle with conditional probability problems and prior studies have shown user accuracy is less than 50%. This study examined user performance when aided with computer-generated Venn and Euler-type diagrams in a non-learning context. Following relational complexity, working memory and mental model theories, this study manipulated problem complexity in diagrams and text-only displays. Partially consistent with the study hypotheses, complex visuals outperformed complex text-only displays and simple text-only displays outperformed complex text only displays. However, a significant interaction between users' spatial ability and the use of diagram displays led to a reversal of performance for low-spatial users in one of the diagram displays. Participants with less spatial ability were significantly impaired in their ability to solve problems with less relational complexity when aided by a diagram.

Keywords: Human-computer interaction, diagrams, Bayesian reasoning, relational complexity, spatial ability, working memory, individual differences, mental models.

1 Objective

Many people struggle with conditional probability problems, which are common in the medical field and in any discipline that needs to communicate and interpret diagnostic evidence [1,2,3]. The research regarding facilitating user performance with computer visuals has yielded conflicting results. The objectives of this study were to a) determine ways of improving user performance with conditional probability problems in non-learning contexts using computer-generated Venn and Euler type diagrams and b) provide a set of design guidelines for designers of systems that have to convey conditional probability problems.

2 Background Literature

According to theories of working memory [4,5] and mental models research [6,7,8,9], visual displays should aid in these tasks as they can make better use of working

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 183–192, 2013.

memory, reduce cognitive load and facilitate mental model construction, thus generating better judgments. The conditional probability problem is a common and difficult quantitative reasoning task. While a significant amount of research has examined why people make errors in conditional probability reasoning problems [10,11,12] very little of it has looked systematically at optimizing people's performance using visuals. Moreover, practically no research has applied theories of working memory to guide facilitation of these kinds of problems.

Counter-intuitively, some research is finding that Venn and Euler type diagrams, which try to clearly depict set inclusions and conditional relationships directly, may actually hurt performance [12,13,14]. Mental models theory [6], relational complexity theory [15] and explanations of working memory limits [16,17] can provide guidance for designing visuals to help people solve these problems.

3 Conditional Probability Problems

Several studies show participants' accuracy on conditional probability problems is low to intermediate, ranging from 6% of the problems correct to 62% depending on the textual format [1,18]. An example of a conditional probability problem, shown below, has been used in several studies [1]:

> *The probability of breast cancer is 1% for women at age forty who participate in routine screening. If a woman has breast cancer, the probability is 80% that she will get a positive mammography. If a woman does not have breast cancer, the probability is 9.6% that she will also get a positive mammography. A woman in this age group had a positive mammography in a routine screening. What is the probability that she actually has breast cancer?*

The normative answer requires applying a form of Bayes' theorem to calculate the posterior probability that a woman with positive results has cancer. The symbols H and $-H$ denote the two hypotheses (cancer and no cancer, respectively) and D denotes the data representing the positive test results. The formula for computing $p(H \mid D)$ is as follows.

$$p(H \mid D) = \frac{p(H)\,p(D \mid H)}{p(H)\,p(D \mid H) + p(-H)\,p(D \mid -H)} = \frac{(.01)(.80)}{(.01)(.80) + (.99)(.096)} = .078$$

Prior studies show that even physicians estimate probabilities incorrectly when presented in text form. Eddy [10] shows that 95% of the physicians queried estimated the probability of p(H | D) in this problem between 70-80% not 7.8%. Other prior research indicates people's accuracy is typically poor for this text-only, probability representation of the problem [1,2,12].

Gigerenzer and Hoffrage [1] have argued that expressing conditional probability problems as natural frequencies rather than as probabilities improves peoples' performance. Natural sampling mimics the process of encountering instances in a population sequentially. Participants should perform better when probability problems are

expressed as natural frequencies. A natural frequency representation has a simpler mathematical form of Bayes' theorem:

$$p(H \mid D) = \frac{d \,\&\, h}{d \,\&\, h + d \,\&\, -h} = \frac{8}{8+95} = 0.078$$

3.1 Use of Visuals in Conditional Probability Problem Solving

A lesser but growing body of research has addressed how to teach people to solve statistical problems [3,11]. Unfortunately, many everyday settings that require statistical reasoning abilities involve people who haven't been taught how to reason with statistics or have failed to retain the associated skills. One common diagram, an Euler diagram, failed to provide significant improvement in either a probability or frequency format [12] and performed worse than iconic displays [13] which depict each item in the frequency of observations as an icon. Iconic displays are superior to Venn diagrams because they better approximate 'actual ecological presentations' [13]. In a study examining deductive reasoning [13], researchers found a disadvantage for Euler diagrams, which performed worse than similar text-only representations.

3.2 Mental Models, Working Memory and Reasoning

Both behavioral and neuropsychological mental models theory research [6,7,9] shows that people reason in these problems not by following linguistic rules, but by means of visuospatial models of the problem structure. While verbal processing and the phonological loop would need to be employed to process the text of these problems and for certain math operations solving the problems requires manipulation of the premises and relations within the problem using visuospatial processing and central executive cognitive resources. Several dual task studies show that the VSSP and CE are involved in problem solving, consistent with mental models theory [8,19]. Mental models need not be visual, and can represent non-spatial relationships, such as kinship, or non-visual precepts in deductive reasoning. Model based reasoning interferes with other spatial tasks, not necessarily with concurrent visual tasks without a spatial component [9]. Due to the overlapping use of the VSSP for both reasoning and visual processing, it is possible that use of visual diagrams might be interfering with performance on reasoning problems.

Cowan [20] defines short term memory as "faculties of the human mind that can hold a limited amount of information in a very accessible state temporarily" and surmises that working memory is used to retain partial results while working out a math problem without paper or "to combine the premises in a lengthy rhetorical argument." Baddeley's [4] influential model of working memory holds that verbal-phonological and visuospatial representations are stored separately and managed and manipulated by the central executive. In Baddeley's model [4], working memory is generally viewed as several components working together [20]. Although early research into working memory capacity showed that people can recall about seven chunks of

information in short-term memory tasks [21] more recent and extensive research is placing a lower limit of about four chunks on the capacity limits in short-term memory [16,17] across both visual and verbal tasks.

3.3 Relational Complexity

Relational complexity is defined as the number of relations that must be processed in parallel, that is, at the same time, to perform a task [15,22] As relational complexity increases, the processing complexity of a cognitive process increases. While a cognitive process may be made up of several steps, the complexity of the process is the measure of the complexity of its most complex task. Process complexity is affected by how many interdependent elements have to be processed in parallel at one time, not how many have to be processed over time. The maximum relational complexity adults can normally process is the quaternary relationship which binds four interactive elements. Planning a correct strategy to solve a problem depends first on representing the complete structure of the problem. Relational theory predicts that people's need to process premises jointly will increase processing load [15]. Mental model construction is constrained by the complexity of the model itself and the working memory capacity, especially the ability for the person to focus attention to the features of the model. In a neurological study [23], researchers confirmed that increasing relational complexity from 0 to 4 in a reasoning problem dramatically increased the percentage of errors and the time in processing, as well as increasing activation in areas of the brain associated with working memory.

4 How Visuals Can Improve Performance

Conditional probability problems typically contain three and can contain as many as six or seven elements that might need to be processed in parallel. Based on relational complexity theory, mental models theory, cognitive load theory and the theories of working memory discussed here, for visuals to reliably help with these kinds of problems, designers need to consider the following:

1. Through perceptual cueing mechanisms such as highlighting and controlling what is visible in the display, restrict the user's focus of attention to the task of relating two independent elements at a time.
2. Use visuals to depict relationships between entities rather than describing the relationship with words.
3. Repeat the process of cueing and processing fewer rather than more elements at a time to help the user make a series of correct inferences.

Using this approach, designers ought to be able to improve user performance in making correct inferences by minimize element interactivity in the display. This encourages users to build partial models based on fewer interacting elements at a time and substitute images for words to depict relationships.

5 Hypotheses

Based on relational complexity theory, working memory theory and mental models research, diagrams perform better than text and displays with lower element interactivity should perform better than those with high element interactivity. This leads to two hypotheses.

- **Hypothesis 1:** Under both high and low levels of relational complexity, users will demonstrate higher performance with diagrams than with text.

Diagrams should help by further relieving working memory by substituting visuals in place of some text (reducing the amount of text processing needed) and by providing visual cues and guidance than can't be equivalently done in a simpler text-only representation. While performance in the diagram or text-only condition is expected to be strongly influenced by complexity, the role of the visual under these two conditions is of interest.

- **Hypothesis 2:** Under both the text and diagram conditions, users will demonstrate higher performance with problems of lower relational complexity than with problems of higher relational complexity.

Since working memory demands will be reduced with low relational complexity displays that have less element interactivity, users will generate more correct answers. Some of the problems with the use of visuals in prior research may be a result of researchers failing to consider the number of interacting elements in conditional probability problems, the overall working memory demands of these problems and the limited working memory resources available.

6 Method

This study recruited 158 participants from a large research university. This study used both students and staff participants with varying ages and backgrounds in order to better approximate real-world populations. This study used a 2X2 factorial design with the two factors being complexity and display type. Complexity has two levels: low (three interacting elements at a time) and high (six interacting elements at a time). Display type has two levels: text-only and diagrams. Because participants were randomly assigned to one of the four groups in the factorial design, this was a between-subjects design. The independent variables were relational complexity (low relational complexity and high relational complexity) and display type (text-only and diagrams). Four treatments were used. Text-only treatments with low and high relational complexity and Venn-type diagram treatments of low and high relational complexity were used and presented via a computer display. Participants expressed their answers in whole numbers and were allowed to round up or down to the next integer. User performance was measured by number of correct answers out of the ten in the battery of problems. Participants had 60 minutes to complete the 10 problems and all

participants completed them within the allotted time. To test for individual differences and their effect on problem performance, participants completed a pre-test survey that collects background information (age, gender, educational level, probability reasoning experience, skill, and whether English is a native language). Because spatial ability may contribute to, if not be the primary source of people's performance with regards to relational complexity [24] and is used for processing spatial diagrams [25], this study measured spatial ability [26] and used it as a covariate.

7 Results

Approximately one-third of the participants (57) were older than 22 and 28 of the participants were not in college. None of the participant background variables were found to be significantly or marginally associated with performance. Performance was measured by the count of correct problems out of the ten problems presented. Table 1 shows the means, standard deviation and sample size for each condition.

Table 1. Performance means

Complexity	Display Type					
	Diagram			*Text*		
	M	SD	n	M	SD	n
Complex	5.38	2.62	40	3.32	2.00	37
Simple	6.39	3.60	44	7.41	3.13	37

Since the hypotheses call for testing for main effect and for interaction with the covariate, a two-way ANCOVA analysis was conducted to test for the influence of complexity and display type and to check for any influence of spatial ability (as measured by the VZ-2 test score). The overall analysis of covariance was highly significant ($F(6,151) = 17.56$, $p<.0001$, $\eta 2=.41$). The analysis of the model effects are shown in Table 2.

Table 2. ANCOVA for Performance

Source	DF	SS	MS	F	Pr > F
Display type	1	11.47	11.47	1.76	0.1860
Complexity	1	236.97	236.97	36.47	<.0001
Spatial ability	1	276.10	276.10	42.50	<.0001
Complexity*Display type	1	62.53	62.53	9.62	0.0023
Display type*Spatial ability	1	75.39	75.39	11.60	0.0008
Complexity*Spatial ability	1	18.30	18.30	2.82	0.0953

This study showed a significant main effect for complexity but not for display type. However, the main effects (complexity and display type) interacted and the covariate (spatial ability) interacted with Display type. The ANCOVA analysis also showed that the slopes of the regression lines for each treatment are not the same, also confirming an interaction between display type and spatial ability.

Pairwise comparisons to test the hypotheses for performance were conducted with one-tailed t-tests to match the hypotheses, adjusted for the differing covariate means and for multiple inferences. Looking at the first hypothesis, H1, participants in this study demonstrated higher levels of performance with diagrams than with text for only the high relational complexity (complex) treatments (5.38 versus 3.32 correct, t(75)=3.29, p=.0034, Cohen's d=.76). Participants did not demonstrated better performance with diagrams than with text in the low relational complexity (simple) condition (6.39 versus 7.41 correct, t(79)=-1.53, p=.7878). Thus H1 is partially confirmed.

Looking at the second hypothesis (H2), participants demonstrated better performance with low relational complexity (simple) displays over high relational complexity (complex) displays for only the text treatments (7.41 versus 3.32 correct, t(72)=6.57, p<.0001, Cohen's d=1.55.) Participants did not demonstrate improved performance with simple displays than with complex ones for the diagram treatments (6.39 versus 5.38 correct, t(82)=1.46, p=.0840). Thus H2 is partially confirmed. Table 3 lists the comparisons for H1 and H2.

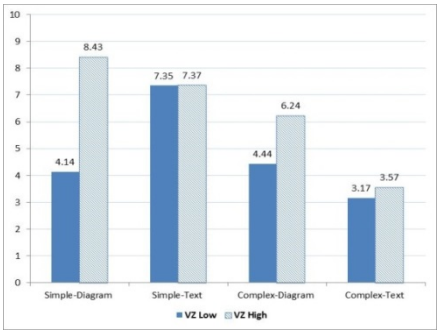
Table 3. Pairwise *comparisons*

	M	SD	M	SD	*t*-test		Cohen's *d*
H1: Diagrams > Text							
3D>3T	6.39	3.60	7.41	3.13	-1.53		
6D>6T	5.38	2.62	3.32	2.00	3.29	**	.76
H2: Low RC > High RC							
3D>6D	6.39	3.60	5.38	2.62	2.07		
3T>6T	7.41	3.13	3.32	2.00	6.54	***	1.55

$^{**}p \le .01$ $^{***}p \le .001$

This study finds partial support for both hypotheses. Using diagrams helped participants in the complex condition. Because diagrams enabled use of visual processing and text processing that conserves working memory and provides cues and guidance not possible in a text display, users performed better. Reducing relational complexity also helps improve performance, but for the text-only condition

In looking at the ANCOVA results, diagram displays and spatial ability are affecting user performance. To understand the nature of the main effects interaction it is helpful to see performance by a split of spatial ability (below). Low spatial participants performed 51% worse (4.14 versus 8.44, N=21 and 23 respectively) on the simple diagram and 28% worse (4.44 versus 6.24, N=19 and 21 respectively) on the complex diagram than the high spatial participants. Additional two-tailed t-tests confirmed this finding for both the simple diagram (3D) displays (t(42)=4.79, p<.0001, Cohen's *d*=1.46) and the 6D



(t(38)=2.31, p<.0264, Cohen's *d*=.74). Also of note, low-spatial and high spatial participants did not differ in their performance on both the simple and complex text-only versions of the problem (p=.5662 and p=.9267 respectively). The diagram displays are impairing low-spatial participants, precipitously for the simple diagram (3D) treatment and significantly for the complex diagram (6D) treatment. The text displays performed equally well for high- and low-spatial participants.

8 Discussion

While overall, reducing relational complexity improves user performance and more complex problems are aided by diagrams, spatial ability is interacting with display type. Diagram displays hurt the performance of participants with less spatial working memory. Spatial ability may be a shared working memory resource serving both non-visual reasoning and perceptual processes. Perceptual processes, being bottom-up processes, may be taking priority and interfering with the reasoning processes for low-spatial users. In terms of Baddeley's model [4], this may mean that the VSSP may be required for non-visual reasoning processes like those used in conditional probability problems as well as for processing an external representation, which in this case is a diagram. Low-spatial users may have less capacity to process or inhibit and restrict bottom-up processing, thus causing visual diagram processing to conflict with spatial working memory needed to construct and validate a correct mental model. While these bottom-up perceptual processes may be the basis of improvement for high-spatial users as it benefits them in offloading cognitive work to the visual stream, these bottom up processes may be crowding out working memory needed for reasoning in the low-spatial user.

Since spatial ability is also associated with other intelligence measures [27,28], one might also expect that low-spatial participants would have reduced performance on the text-only version of the problem due to presumed deficits in other non-visual cognitive capacities. This study failed to identify that effect. Low and high spatial users performed similarly on text-only versions of the problem. Since this study relied exclusively on the computer presentation of each problem by not allowing any secondary tools such as notepaper to interact with, participants had no choice but to interact with the computer display to solve the problem. This study suggests that capacity limits in spatial working memory may be the culprit for the diminished diagram performance for low-spatial users.

To help users with conditional probability problems in everyday, non-learning contexts, this study recommends the following guidelines:

1. Reduce relational complexity. With the exception of the simple diagram display for low-spatial users, simpler problems helped users produce more correct answers.
2. For low-spatial users, use simpler, text-only displays. These displays will facilitate performance better than diagrams will.
3. For high-spatial users, Venn and Euler-type diagrams can be safely used and can improve performance.

4. Use natural frequencies. All of the problems used in this study used natural frequencies and the best performing treatment, the simpler text-only displays, had a mean of 7.51 out of 10 correct. If all components of the conditional probability problem must be displayed at once, the diagrammatic display will provide the best performance overall.
5. If spatial ability measures are available for the target audience, designers can personalize the display. Simpler diagram displays can be used with high-spatial users and simpler text-only displays can be used with low-spatial users.

9 Conclusion

This study has further clarified how diagrams can (or cannot) aid in solving conditional probability problems. Reductions of complexity can improve performance, especially with text displays. Diagrams can help, but they can hurt too. Individual differences in spatial ability matter when processing Venn and Euler-type diagrams. The impairment with diagrams for people with lower spatial ability could be the source of some of the conflicting findings on the use of diagrams to facilitate conditional problem. The extensive collection of research into people's performance in solving conditional probabilities and Bayesian reasoning has not yet applied concepts and frameworks from cognitive psychology like relational complexity and theories of working memory. The recent research into mental models and the neural correlates of working memory and reasoning can provide linkages between the functional and physical descriptions of the mind. These frameworks can be applied to improve our understanding of user performance problems with computer diagrams without having to refer to less testable constructs such as 'frequency coding in the mind' posited by frequentist interpretations of performance [18].

References

1. Gigerenzer, G., Hoffrage, U.: How to improve Bayesian reasoning without instruction: frequency formats. Psychological Review 102, 684–701 (1995)
2. Kahnamen, D., Lovallo, D.: Timid choices and bold forecasts: a cognitive perspective on risk taking. Management Science 39(1), 17–31 (1993)
3. Sedlmeier, P.: How to improve statistical thinking: Choose the task wisely and learn by doing. Instructional Science 28, 227–262 (2000)
4. Baddeley, A.: Working memory. Clarendon Press, Oxford (1986)
5. Miyake, A., Shah, P.: An Introduction. In: Miyake, A., Shah, P. (eds.) Models of Working Memory. Cambridge University Press (1999)
6. Johnson-Laird, P.N., Legrenzi, P., Girotto, V., Legrenzi, M.S.: Naive probability: A mental model theory of extensional reasoning. Psychological Review 106, 62–88 (1999)
7. Fangmeier, T., Knauff, M., Ruff, C., Sloutsky, V.: fMRI evidence for a three-stage model of deductive reasoning. Journal of Cognitive Neuroscience 18(3), 320–334 (2006)
8. Knauff, M., Jola, C., Strube, G.: Spatial reasoning: No need for visual information. In: Montello, D.R. (ed.) COSIT 2001. LNCS, vol. 2205, pp. 447–457. Springer, Heidelberg (2001)

9. Ruff, C.C., Knauff, M., Fangmeier, T., Spreer, J.: Reasoning and working memory: Common and distinct neuronal processes. Neuropsychologia 41, 1241–1253 (2003)
10. Eddy, D.M.: Probabilistic reasoning in clinical medicine: Problems and opportunities. In: Kahneman, D., Slovic, P., Tversky, A. (eds.) Judgment under Uncertainty: Heuristics and Biases, pp. 249–267. Cambridge University Press, Cambridge (1982)
11. Gigerenzer, G., Hoffrage, U.: Overcoming difficulties in Bayesian Reasoning: A reply to Lewis & Keren and Mellers & McGraw. Psychological Review 106, 425–430 (1999)
12. Sloman, S.A., Over, D., Slovak, L., Stibel, J.M.: Frequency illusions and other fallacies. Organizational Behavior and Human Decision Processes 91, 296–301 (2003)
13. Brase, G.L.: Pictorial representations in statistical reasoning. Applied Cognitive Psychology (2008), doi:10.1002/acp.1460
14. Calvillo, D.P., DeLeeuw, K.E., Revlin, R.: Deduction with Euler Circles: Diagrams That Hurt. In: Barker-Plummer, D., Cox, R., Swoboda, N. (eds.) Diagrams 2006. LNCS (LNAI), vol. 4045, pp. 199–203. Springer, Heidelberg (2006)
15. Halford, G.S., Wilson, W.H., Phillips, S.: Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. Behavioral and Brain Sciences (21), 803–865 (1998)
16. Cowan, N.: The magical number 4 in short-term memory: A reconsideration of mental storage capacity. Behavioral and Brain Sciences 24, 87–185 (2000)
17. Cowan, N., Elliott, E.M., Saults, J.S., Morey, C.C., Mattox, S., Hismjatullina, A., Conway, A.R.: On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. Cognitive Psychology 51, 42–100 (2005)
18. Brase, G.L., Cosmides, L., Tooby, J.: Individuation, counting, and statistical inference: the role of frequency and whole-object representations in judgment under uncertainty. Journal of Experimental Psychology: General 127(1), 3–21 (1998)
19. Klauer, K.C., Stegmaier, R., Meiser, T.: Working memory involvement in propositional and spatial reasoning. Thinking and Reasoning 3(1), 9–47 (1997)
20. Cowan, N.: What are the differences between long-term, short-term, and working memory? In: Sossin, W.S., Lacaille, J.-C., Castellucci, V.F., Belleville, S. (eds.) Progress in Brain Research, vol. 169. Elsevier B.V. (2008)
21. Miller, G.A.: The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review 63, 81–97 (1956)
22. Halford, G.S., Wilson, W.H., Phillips, S.: Relational knowledge: The foundation of higher cognition. Trends in Cognitive Sciences 14(11), 497–505 (2010)
23. Kroger, J.K., Sabb, F.W., Fales, C.I., Bookheimer, S.Y., Cohen, M.S., Holyoak, K.J.: Recruitment of anterior dorsolateral prefrontal cortex in human reasoning: a parametric study of relational complexity. Cerebral Cortex 12, 477–485 (2002)
24. Halford, G.S., Cowan, N., Andrews, G.: Separating cognitive capacity from knowledge: A new hypothesis. Trends in Cognitive Science 11(6), 236–242 (2007)
25. Hegarty, M.: Individual differences in use of diagrams as memory in mechanical reasoning. Learning and Individual Differences 9(1), 19–42 (1997)
26. Ekstrom, R.B., French, J.W., Harman, H.H.: Manual for kit of factor-referenced cognitive tests. Educational Testing Service, Princeton (1976)
27. Hegarty, M., Waller, D.: Individual differences in spatial ability. In: Shah, P., Miyake, A. (eds.) The Cambridge Handbook of Visuospatial Thinking, pp. 121–169. Cambridge University Press, Cambridge (2005)
28. Miyake, A., Friedman, N.P., Rettinger, D.A., Shah, P., Hegarty, M.: How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. Journal of Experimental Psychology: General 130(4), 621–640 (2001)

Establishing a Cognitive Map of Public Place for Blind and Visual Impaired by Using IVEO Hands-On Learning System

Qing-Wen Lin[1], Sheue-Ling Hwang[1], and Jan-Li Wang[2]

[1] National Tsing Hua University, Hsinchu, Taiwan 300,
slhwang@ie.nthu.edu.tw, overalldream@gmail.com
[2] National Center for High-Performance Computing, NARL, Hsinchu, Taiwan 300
wangjanli@nchc.narl.org.tw

Abstract. In recent years, there has been an escalation of orientation and way-finding technologies and systems for the blind visually impaired people. The purpose of this study was to help visual impaired to access and explore spatial information by establishing a cognitive map of public place using IVEO hands-on learning system and tactile direction map. The IVEO hands-on learning system combines touch, sound, and sight, using touch and sound learning modalities to help the blind and visually impaired to learn faster and retain information longer, meanwhile making learning more interesting and interactive. Subjects were divided into four groups in terms of route familiarity and IVEO Aids of exposure, with an equal number of subjects (N=4) in each group. Then, the subject performed an actual Way-finding test to get back and forth between Ximending and TDTB (Taiwan Digital Talking Books association). The dependent variables were total time, number of miss route, and number of request. The results revealed a significant effect of IVEO Aids, and the subjects could perform better through IVEO Aids. This research provided a learning method of orientation and mobility for the blind and visually impaired, and assisted them to be familiar with the route safer and faster.

Keywords: Learning System, audio-tactile maps, cognitive map, Way-finding, blind people.

1 Introduction

1.1 Background and Motivation

According to the statistics by the Ministry of Interior, Taiwan, the legal population of the visually impaired increased from 44889 persons in 2002 to 56589 persons in 2012 in Taiwan. The population of the blind and the visually impaired has been increasing sharply over the past 10years. (Taiwan Ministry of the Interior, 2012).

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 193–198, 2013.

In an alien environment, the sighted adult usually need a map to find spatial information while a tactile direction map is required for the blind and visually impaired (Levine, 1982; O'Neill, 1999; Klippel, Freksa & Winter, 2006).

The blind and visually impaired can successfully use verbally presented route information as well as survey information in constructing cognitive maps of environments, even if they are not as efficient as normally sighted people (Steyvers and Kooijman, 2008).Thus audio-tactile maps, which allow a multi-modal co-comprehension of tactilely perceived and verbally presented information, are a promising way to diminish some restrictions of tactile maps.

1.2 Objectives

The IVEO hands-on learning system combines touch, sound, and sight, using touch and sound learning modalities to help the blind and visually impaired to learn faster and retain information longer, meanwhile making learning more interesting and interactive (Gardner, 2006). We propose assistive technology system to aid in way-finding based on IVEO hands-on learning system. The participants who learned the route by using IVEO Aids and tactile direction map showed a more accurate level of confidence in retrieving the spatial information necessary to complete the way-finding. The IVEO hands-on learning system is shown in Figure 1.

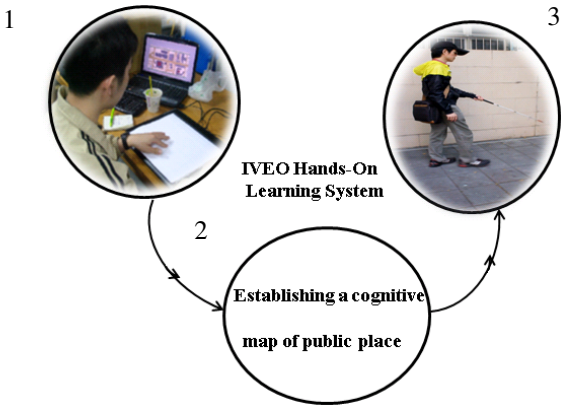


Fig. 1. The IVEO hands-on learning system

Tactile maps can also be an extremely effective tool for representing spatial information for the orientation and mobility (O&M) student (Bentzen, 1997). Blades et al. (2002) found that the performance of visually impaired people can be enhanced by the strategies used for learning a new route, e.g. by modeling a map of the route itself. However, many orientation and mobility (O&M) counselors dedicate a great portion of their lesson preparation time to the founding of detailed tactile maps to be used

only once or twice by the student. An IVEO creator software and View-Plus Embosser for producing highly detailed, "one-off" maps would be of great advantage to these counselors and their students, freeing the counselor's time for actual O&M counselor, instead of the difficult task of tactile cartography.

2 Research Method

2.1 IVEO Hands-On Learning System Interface Design

Automatic tactile-audio map creation is the key component of our proposed interactive solution. The step takes a single map image as input and generates output image for easy tactile reproduction and an SVG file which combines graphical contents with audio annotations. We applied the mainstream graphics software, such as Google Sketch Up to design the map image, and interviews and questionnaire were designed to identify the requirements that support the blind and visually impaired for Way-finding activities. Moreover, use IVEO creator software to input the title and content to customize tactile map. In addition, View-Plus Embosser was used to print the tactile map and place it on the IVEO touchpad so that the users can press different areas of the map to hear the titles and content spoken back to them. The Processing flow of the IVEO System Interface Design in Figure 2.

2.2 Experimental Design

The primary criterion for subjects selection was that they are the blind visually impaired people. Moreover, they have experience to walk alone in the streets. The subject was asked to fill out a SDQS/sense of direction questionnaire-short form which elicited information concerning his cognitive map, sense of direction, spatial memory, familiarity and uniform into four groups accordance with scores. The independent variables were route familiarity (low or high) and IVEO Aids of exposure (with or without). The low familiarity was defined as one practice of walking from Ximending to TDTB before proceeding Way-finding test, and the high familiarity was defined as three times practices of walking from Ximending to TDTB.

A total of 16 visually impaired participated in this experiment, and the subject was assigned to one of four groups in combination of experimental variables, familiarity and IVEO Aids of exposure, with an equal number of subjects (N=4) in each group. R1 group is low route familiarity and without IVEO Aids, R2 group is high route familiarity and without IVEO Aids, R3 group is low route familiarity and with IVEO Aids, R4 groups is high route familiarity and with IVEO Aids. The dependent variables were total walking time, number of miss routes, and number of requests.

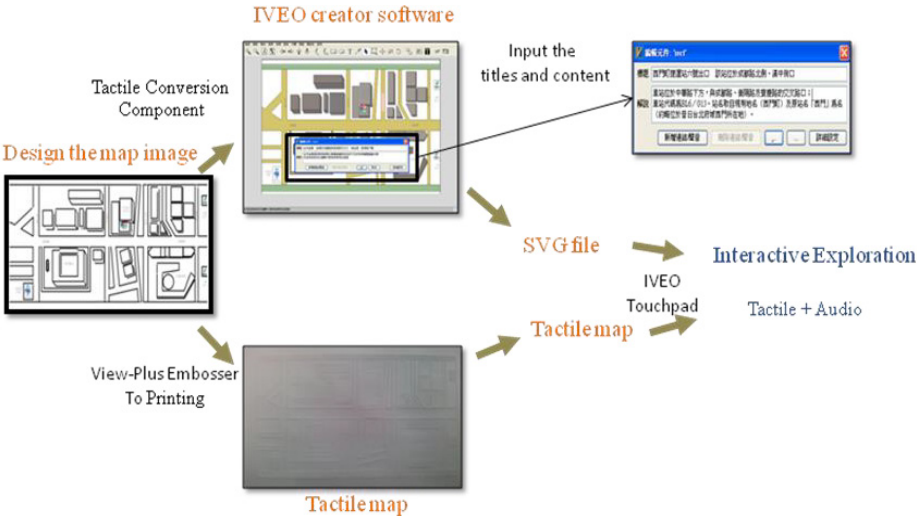


Fig. 2. Processing flow of the IVEO System Interface Design

2.3 Experimental Procedure

The experiment took around 3 hours, following steps:

1. The subject was asked to fill out a SDQS/sense of direction questionnaire-short form.
2. The subject performed a Way-finding test from Ximending to TDTB (Taiwan Digital Talking Books association). Data collected primarily by two experimenters. One was a cameraman who recorded all the scenes with a video camera along the way. Another logger registered behaviors of the subjects and record his/her specific activities.
3. Following the test, subjects were interviewed and asked to fill out NASA-TLX questionnaire and interviewed for approximately half an hour about their mental workload.

3 Experimental Results

3.1 Number of Missed Routes

Analyses of variance were used to detect significant difference among variables. Table 1 showed two-way ANOVA results of number of missed route. There was a significant effect of IVEO Aids on number of missed route ($p<0.05$), but the effect of route familiarity was not significant. It means that with the IVEO Aids in actual Way-finding test, the Number of Missed routes of subjects would decrease significantly.

Table 1. Two-way ANOVA on number of miss routes

Source	DF	SS	MS	F	P
IVEO Aids	1	45.563	45.563	8.64	0.012*
Route Familiarity	1	22.563	22.563	4.28	0.061
Interaction	1	5.063	5.063	0.96	0.346
Error	12	63.250	5.2708		
Total	15	136.438			
S = 2.296 R-Sq = 53.64% R-Sq(adj) = 42.05%					

*p<.05.

3.2 Number of Requests

Likewise, number of requests were calculated during the experiment. Table 2 showed two-way ANOVA results on number of request. There was a significant effect of IVEO Aids on number of requests ($p<0.05$), and the effect of route familiarity was also significant ($p<0.01$).The route familiarity and IVEO Aids for the number of requests is of decisive importance in actual Way-finding test.

Table 2. Two-way ANOVA on number of requests

Source	DF	SS	MS	F	P
IVEO Aids	1	9	9	8.64	0.012*
Route Familiarity	1	12.25	12.25	11.76	0.005**
Interaction	1	4	4	3.84	0.074
Error	12	12.5	1.0417		
Total	15	37.75			
S = 1.021 R-Sq = 66.89% R-Sq(adj) = 58.61%					

*p<.05 **p<.01.

3.3 NASA-TLX

Subjects were asked to fill out a NASA-TLX questionnaire for measuring the mental workload. According to the score of each item, the differences of mental workload in these subjects were evaluated. Table 3 showed two-way ANOVA results of NASA-TLX. The result could help us to evaluate the mental workload for cumulative learning effect. There was a significant effect of IVEO Aids on NASA-TLX score($p<0.01$) that mental workload of subjects would decrease significantly by IVEO Aids, and the effect of route familiarity was also significant ($p<0.05$).

Table 3. Two-way ANOVA on score of NASA-TLX

Source	DF	SS	MS	F	P
IVEO	1	2620.08	2620.08	46.69	0.000**
Route Familiarity	1	369.73	369.73	6.59	0.025*
Interaction	1	240.04	240.04	4.28	0.061
Error	12	673.47	56.12		
Total	15	3903.32			
SS = 7.491 R-Sq = 82.75% R-Sq(adj) = 78.43%					

*p<.05 **p<.01.

4 Conclusion

We can find significant effect of using IVEO Hands-on learning system for the blind and visually impaired. Thus subjects could perform better through IVEO Aids in terms of IVEO Aids and route familiarity. This research provided a learning method of orientation and mobility for the blind and visually impaired with orientation and mobility (O&M) counselors, and assisted them to be familiar with the route safer and faster.

Acknowledgments. Partial of this work was sponsored by the National Science Council(NSC), Taiwan, R.O.C. through project NSC 101-2511-S492-001 and the work was enhanced by the gracious assistance of Ludy Lee who is Secretary General of Taiwan Digital Talking Books association (TDTB).The authors like to express their appreciation to all of the members involved in this project.

References

1. Gardner, J.A., Bulatov, V.L.: Scientific Diagrams Made Easy with IVEO^{TM}. In: Miesenberger, K., Klaus, J., Zagler, W.L., Karshmer, A.I. (eds.) ICCHP 2006. LNCS, vol. 4061, pp. 1243–1250. Springer, Heidelberg (2006)
2. Klippel, A., Freksa, C., Winter, S.: You-Are-Here Maps in Emergencies — The Danger of Getting Lost. Journal of Spatial Science 51(15), 117–131 (2006)
3. Kobayashi, K., Lee, R., Sumiya, K.: Lost again on the way?: Measuring Human Map-Reading Ability. DBSJ Journal 8(4) (2010)
4. Parente, P., Bishop, G.: BATS: the blind audio tactile mapping system. In: Proc. of the ACM Southeast Regional Conference (2003)
5. Taiwan's Ministry of the Interior, statistics the population of blind and the visually impaired persons (2012), `http://sowf.moi.gov.tw/stat/year/list.htm` (retrieved)

An Application of the Ballistic Movement Method for Evaluating Computer Mice

Ray F. Lin, Ching-Wen Chung, Yi-Chien Tsai, and Chi-Yu Huang

Department of Industrial Engineering & Management, Yuan Ze University,
135 Yuan-Tung Road, Chungli, Taoyuan, Taiwan 32003
juifeng@saturn.yzu.edu.tw

Abstract. Lin and Drury [1] validated two ballistic movement models and suggested these models could have certain benefits for evaluating input devices. Hence, the study aimed at (1) validating the application of ballistic movement models for evaluating computer mice and (2) demonstrating the superiority of this method, compared to Fitts' law. In a two-stage experiment, four participants used six types of computer mice to execute Fitts-type aiming movements and ballistic movements, sequentially. The measured data were analyzed by Fitts' law and the ballistic movement models, respectively. The comparison of the results obtained by the two methods showed that (1) ballistic movement models can well fit the measured data and (2) the ballistic movement method can provide independent performance information of "speed" and "accuracy" that is not available by applying Fitts' law. This study demonstrated an alternative method for evaluating computer mice.

Keywords: ballistic movement method, computer mouse, input device, human movements, Fitts' law.

1 Introduction

A computer mouse is one of the most effective and efficient input devices while interacting with computers. To enhance the performance of this human-machine system, researchers and designers make efforts to develop a variety of computer mice. In the development procedures, it is necessary to evaluate developing mice according to several criteria, such as production cost, durability, and usability. Most importantly, task performance is an essential criterion for assessing computer mice.

To evaluate the task performance of computer mice, Fitts' law [2] has been accepted as a useful approach, which is easy to apply and widely used. However, Fitts' law has limitations of lacking theoretical support and confounding performance of movement speed and accuracy. Lin and Drury [1] suggested a new evaluation method, called the "ballistic movement method", for evaluating input devices. They stated that the method could provide independent performance of movement speed and accuracy. However, the method has not been tested in evaluating computer mice. Hence, it is the goal of this study to test the application of the ballistic movement method, compared to Fitts' law.

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 199–208, 2013.

1.1 Fitts' Law

As shown in Equations 1 and 2, Fitts' law [2] predicts that the movement time (*MT*) required to execute a Fitts-type aiming movement is linearly related to the index of difficulty (*ID*) of that movement, defined as the dyadic logarithm of the quotient of amplitude of the movement and target width (Equation 2).

$$MT = a + b \times ID \tag{1}$$

$$ID = log_2 \frac{2A}{W} \tag{2}$$

where *A* is movement amplitude, *W* is target width, *a* and *b* are experimentally determined constants.

To compare two computer mice, for example, the practitioner asks participants to perform Fitts-type aiming movements manipulated at several values of index of difficulty. Fitt's law, hence, is able to describe the tradeoff relationships between the movement accuracy and speed by using two computer mice. By regressing measured *MT* data on to manipulated *ID* values for two computer mice and then drawing two regression lines on a two dimensional graph, researchers can easily compare the performance of the two computer mice. Because Fitts-type aiming movements are easily tested and the measured data can be well predict by Fitts' law, Fitts' law thus become one of the most popular evaluation methods in the domains of Human Factors and Human Computer Interaction.

Although Fitts' law is easy to apply, as mentioned in Lin and Drury [1], Fitts' law has two limitations. First, Fitts' law only allows practitioners obtain the performance information that is confounded with the two motor properties: speed and accuracy. A Fitts-type aiming movement that takes a longer movement time could result from lower motor speed, lower motor accuracy, or a combination of both. However, Fitts' law has difficulty discriminating the extent to which the two motor properties contribute to the overall movement time. Second, the application of Fitts' law has theoretical issues while reporting the throughput values and generalizing the measured results to other conditions [see 3 for detials].

1.2 Ballistic Movement Method

In recent year, the general model proposed by Lin, Drury, Karwan and Paquet [4] and Lin and Drury [5] indicates that a Fitts-type aiming movement is composed of ballistic movements, which are basic movement unit. The movement time and the endpoint variability of a ballistic movement are two essential factors that directly affect the speed and accuracy of a Fitts-type aiming movement. Lin and Drury [1] further tested two ballistic movement models for describing how these two properties are associated to ballistic movement distance.

Ballistic movement time represents the required time for performing a ballistic movement. In an experiment in which participants performed hand control movements on a drawing tablet, Lin and Drury [1] verified that Equation 6, proposed by Gan & Hoffmann [6], can effectively describe and predict the relationship between

ballistic movement time ($t_{ballistic}$) and the squared root of ballistic movement distance ($\sqrt{d_u}$).

$$t_{ballistic} = i + j \times \sqrt{d_u} \tag{3}$$

where *i* and *j* are experimentally determined constants.

Ballistic movement variability describes the endpoint variability of a ballistic movement. Because of certain noise existing in human motor control mechanism, the ultimate endpoint of a ballistic movement may not exactly end at the anticipative aimed point [7, 8]. No matter endpoint errors are measured in the movement direction or perpendicular to the movement direction, Lin and Drury [1] found that the probability of endpoint location formed a normal distribution around the aimed point. In order to predict two directions of endpoint variability, Lin and Drury [1] verified the application of Equation 7, originally developed by Howarth, Beggs and Bowden [9].

$$\sigma^2 = e + f \times {d_u}^2 \tag{4}$$

where *e* and *f* are experimentally determined constants. As shown in the equation, the endpoint variability is linearly related to the square of movement distance.

The two ballistic movement models (i.e., Equations 3 and 4) have been tested in several conditions by Lin and his colleagues [1, 10, 11]. Lin and Drury [1] originally verified the two models by asking participations to perform ballistic movements using a drawing tablet. Further, the models were tested in executing three-dimensional hand movements [10] and evaluating age effects on touchscreen usage [11]. With the ballistic movement method, Lin and Ho [10] were able to evaluate three-dimensional hand movements in a detail manner in which the movement speed and accuracy could be evaluated independently. Six right-handed graduated students participated in their study. In terms of movement speed, some participants performed differently while using left hand, compared to right hand, but some kept the same no matter which hand was used. In terms of movement accuracy, however, all the participants had higher accuracy in any of the three dimensions while using their right hands. Lin, et al. [11] attempted to use the ballistic movement method for evaluating ageing difference while using a touchscreen. Although no significant difference were found on ageing effects by analysis of variance, the ballistic movement models showed that the older participants performed slower movements, but had greater movement accuracy, compared to the young participants.

1.3 Research Objective

Although Fitts' law and Fitts-type aiming movements have been widely used as an evaluation method, the method has theoretical issues and only provides confounded results of movement speed and accuracy. Nonetheless, ballistic movements are essential movement unit to construct an aiming movement. The performance of executing ballistic movements is associated to the performance of aiming movements. The two ballistic movement models introduced above separate the movement performance into "speed" and "accuracy", independently. Hence, this study aimed at verifying the

ballistic movement models as a new evaluation method for computer mice. This new method was expected with superiority and more effectiveness, compared to Fitts' law.

2 Method

2.1 Participants and Equipment

Two female and two male undergraduate students, aged from 19 to 20 years, participated in this study. They were all right-handed with normal or corrected-to-normal vision.

Fig. 1. Six computer mice tested in this study

Experimental apparatus included a personal computer (PC), six types of computer mice (see Fig. 1), and two self-developed programs, comprising the Fitts-type aiming movement program and the ballistic movement program. The PC ran Visual Basic 6.0 using the developed programs that both displayed the experimental tasks and measured task performance.

2.2 Experimental Procedures

After informed consent procedures, the four participants utilized six computer mice to performed both Fitts-type aiming movements and ballistic movements. To reduce training effect, the participants had sufficient time (about 2-3 hours) to practice all the computer mice and the experimental programs. Because Fitts-type aiming movements were relatively easier to perform, all the participants executed Fitts-type aiming movements in the first stage of experiment and then executed ballistic movements in the second stage of experiment. To reduce the fatigue effect, the participants finished only a measurement that tested one computer mouse in a half-day.

To perform Fitts-type aiming movements, as shown in Fig. 2(a), the participants drew a line from the starting point to end within the target line. The movements were all performed from the left to the right. The participants were asked to complete every trial as fast as possible. If the cursor was moved over the target line, the participants needed to immediately change the moving direction and continue the movement. By clicking the cursor on any location of the screen, the participants could continue on the next trail.

To perform ballistic movements, as shown in Fig. 2(b), the participants drew a line from a starting point to the center of the cross target with a certain distance (d_u). The movements were all performed from left to right as well. The tasks started by pressing down the stylus cursor on the starting point and then moving toward the cross target. Once the cursor was moved away from the starting point, the cursor and the cross target disappeared and the movement time started to record. When the movement stopped, the cross target and the endpoint of that movement were immediately displayed on the screen. Similar to Fitts-type aiming movements, participants could continue on the next trial by clicking the cursor on the screen.

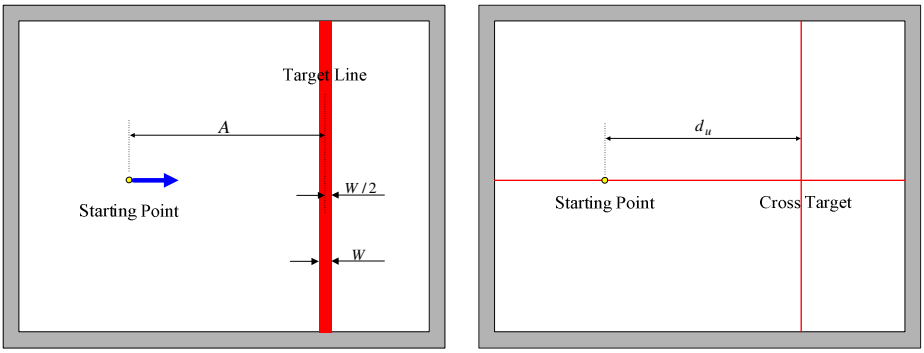


Fig. 2. The execution of Fitts-type aiming movements (a) and ballistic movements (b)

In order to reduce kinesthetic feedback caused by repetitive movements from the same starting point, there were four different starting points decided by a horizontal displacement (200 pixels) and a vertical displacements (200 pixels) in both the experiments.

2.3 Experimental Variables

In the experiment of Fitts-type aiming movement, the independent variables were Computer Mouse and *ID* and the only dependent variable were Movement Time. The four *ID* values were 2, 3, 4, and 5 bits determined by the modified Fitts' law [12, 13]. The experimental program of Fitts-type aiming movement replicated each *ID* movements 24 times, resulting in a total of 96 trials. All the trials were randomly conducted by each participant, taking about 15 minutes to finish.

In the experiment of ballistic movement, the independent variables were Computer Mouse and Movement Distance and the dependent variables were Movement Time and the movement errors measured in the movement direction (X error) and perpendicular to movement direction (Y error). The errors consisted of constant error and variable error. To analyze whether the independent variables had significant effects on these two types of errors, five replications of each experimental combination were utilized to calculate the constant error (measured by mean) and the variable error (measured by variance). However, only the results of variable error are discussed in

this article. The 14 values of movement distance (d_u) were 8, 17, 32, 53, 80, 113, 152, 197, 248, 305, 368, 437, 512, and 593 pixels. The experimental program of ballistic movement replicated each movement distance 20 times, resulting in a total of 280 trials. All the trials were randomly conducted by each participant, taking about 20 minutes to finish.

3 Results

3.1 Analysis of Variance

Analysis of variance was first performed to test the effects of independent variables on dependent variables in the two experiments. In the experiment of Fitts-type aiming movement, the main effects and the interaction effect of Computer Mouse and *ID* had highly significant effects on Movement Time (all the p values < 0.001). In the experiment of ballistic movement, the main effects and the interaction effect of Mouse and Movement Distance also had highly significant effects on Movement Time and two directions of variable errors (all the p values < 0.001). Since all the main effects in the two experiments had significant effects on the dependent variables, the applications and comparison of Fitts' law and ballistic movement models were performed.

3.2 Results Obtained by Applying Fitts' Law

The means of aiming movement time (*MT*) of the six computer mouse were regressed on to the index of difficulty (*ID*) to give six regression lines. Fitts' law accounted for 98.8 % variance on average and at least 97.2 % variance of the data of movement time. The regression lines of the six computer mice are shown in Fig. 3, which shows

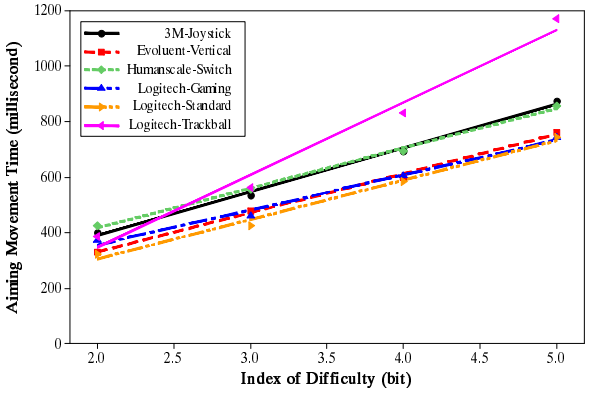


Fig. 3. Fittings of Fitts' law on aiming movement time by using six computer mice

good model fittings. In general, the movement times required to execute aiming movements are close while using the computer mice of Logetich-Standard, Evoluent-Vertical, and Logitecfh-Gaming, which had the best performance compared to the rest. However, the computer mouse of 3M-Joystick required the longest movement time for executing aiming movements, especially when the value of index of difficulty were greater than three.

3.3 Results Obtained by Applying the Ballistic Movement Method

The means of ballistic movement time ($t_{balllitic}$) of the six computer mouse were regressed on to the square root of ballistic movement distance ($\sqrt{d_u}$) to give the slopes and intercepts. The model fitted the data very well. It accounted for 96.2 % variance on average and at least 85.6 % variance of data of ballistic movement time. The regression lines of the six computer mice are shown in Fig. 4, which also shows good model fittings. As shown in the figure, the movement times required to execute ballistic movement were close by using the computer mice of Evoluent-Vertical and Logitech-Trackball, which resulted in the fastest movements. The ballistic movements were also fast by using the computer mice of Humanscale-Switch and Logitech-Gaming. However, the computer mouse of 3M-Joystick resulted in the slowest ballistic movements.

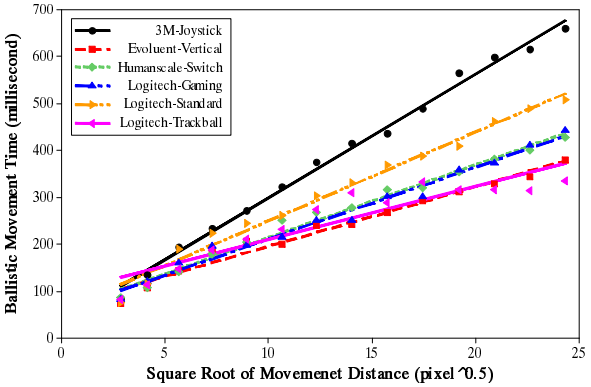


Fig. 4. Fittings of the ballistic movement time model by using the six computer mice

The two error variances, calculated from the raw data for each distance, were regressed on to ${d_u}^2$ to give the slopes, intercepts, and $\boldsymbol{r^2}$ values. For X-variable error (${\sigma_X}^2$), the ballistic movement time model accounted for 93.0 % variance on average and at least 85.5 % variance of data. For Y-variable error (${\sigma_Y}^2$), the ballistic movement time model accounted for 90.4 % variance on average and at least 76.1 % variance of data. The regression lines of X-variable error and Y-variable error are shown in Fig. 5, where X-variable error is about ten times larger than Y-variable error. In terms of X-variable error (measured in the movement direction), the six computer mice listed in descending order by endpoint variance were Logitech-Trackball,

Humanscale-Switch, Evoluent-Vertical, 3M-Joystick, Logitech-Gamming, and Logitech-Standard. The computer mouse of Logitech-Standard resulted in the smallest X-endpoint variance, and the computer mouse of Logitech-Trackball resulted in the greatest X-endpoint variance. In terms of Y-variable error (measured perpendicular to the movement direction), the six computer mice listed in descending order by endpoint variance were 3M-Joystick, Logitech-Trackball, Humanscale-Switch, Logitech-Gamming, Evoluent-Vertical, and Logitech-Standard. Again, the computer mouse of Logitech-Standard resulted in the smallest Y-endpoint variance. The computer mouse of Logitech-Trackball also had great Y-endpoint variance, but the computer mouse of 3M-Joystck had the greatest Y-endpoint variance.

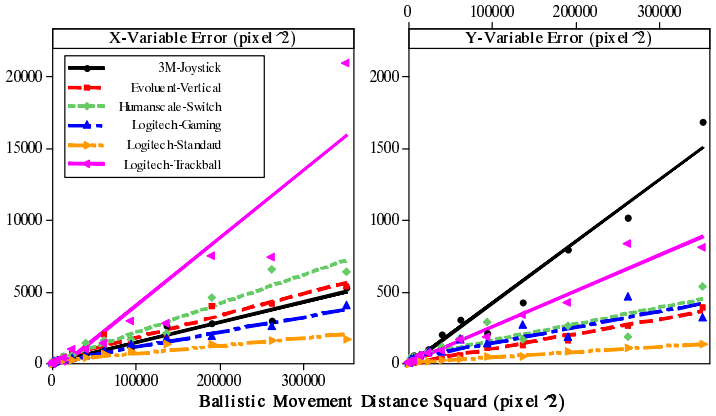


Fig. 5. Fittings of the ballistic movement variability model on the error variance by using the six computer mice

4 Discussion

This study first showed that the ballistic movement model, as Fitts' law, have good predictions of the measured data. Fitts' law, as expected, well predicts the relationship between the movement time and the index of difficulty while using each computer mouse to execute Fitts-type aiming movements. The two ballistic movement models also well predict the relationships between the measured data of ballistic movement time, X-variable error, Y-variable error and the movement distance. Although the fittings of the models were not as good as that of Fitts' law, the models accounted for 96.2 %, 93.0 %, and 90.4 % of three types of data variance, respectively.

While the evaluation results of the two methods were compared, this study then showed the superiority of the ballistic movement method. By applying Fitts' law, we can easily obtain the movement times required by each computer mouse to finish aiming movements set at a certain *ID* value by comparing the linear regression lines obtained by Fitts' law. However, as mentioned previously, the movement time obtained by Fitts' law is confounded by two motor properties of speed and accuracy. Two computer mice may require same aiming movement times to operate, but they

may possess different combinations of speed and accuracy properties. By applying the ballistic movement method, we can further obtain independent performance of speed and accuracy properties of each computer mice by comparing the linear regression lines obtained by the ballistic movement models. This superiority is fully supported by our results. For example, by applying Fitts' law, we only know that the computer mice of Logitech-Standard and Evoluent-vertical required close movement times to operate Fitts-type aiming movements (see Fig. 3). However, the application of ballistic movement method showed these two computer mice possess different combination of speed and accuracy properties. The computer mouse of Logitech-Standard had better performance of accuracy (see Fig. 5), but had poor performance of speed than the computer mouse of Evoluent-vertical (see Fig. 4). Moreover, by applying Fitts' law, we know that the computer mouse of Logitech-Standard required shorter movement times compared to the computer mice of Evoluent-Vertical and Logitech-Gaming. However, Fitts' law provides no information whether the good performance of Logitech-Standard is due to its speed property or accuracy property. By using the ballistic movement method, as shown in Fig. 4 and Fig. 5, we realize that the reason why the mouse of Logitech-Standard required the shortest movement time was not due to its speed performance, but due to its highest accuracy performance. Finally, by applying the ballistic movement method, we know the reason why the computer mouse of Logitech-Trackball required the longest movement time was not due to its speed property, but due to its accuracy property that was the worst among the six computer mice.

Future research should validate the ballistic movement method in a comprehensive way by recruiting adequate participants and controlling driver difference of computer mice. Furthermore, the independent performance of speed and accuracy could help designers focus on specific properties of computer mouse for improvement.

5 Conclusions

This study attempted to validate the two ballistic movement models proposed by Lin and Drury [1] as an alternative method for evaluating the task performance of computer mice. To this end, six types of computer mice were tested by simultaneously applying both Fitts' law and the ballistic movement method. By comparing the results obtained by the two methods, we showed the superiority of the ballistic movement method, compared to Fitts' law. Fitts' law only helps determine the time required performing Fitts-type aiming movements for each computer mouse, whereas the ballistic movement method provides independent performance information about speed and accuracy, which could help effectively improve the design of computer mouse in the future research.

Acknowledgment. We would like to acknowledge the grant support from Taiwan National Science Council (NSC101-2221-E-155-006) for funding the paper submission and presentation.

References

1. Lin, J.-F., Drury, C.G.: Verification of two models of ballistic movements. In: Jacko, J.A. (ed.) Human-Computer Interaction, Part II, HCII 2011. LNCS, vol. 6762, pp. 275–284. Springer, Heidelberg (2011)
2. Fitts, P.M.: The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology 47, 381–391 (1954)
3. Zhai, S.: Characterizing computer input with Fitts' law parameters - the information and non-information aspects of pointing. International Journal of Human-Computer Studies 61, 791–809 (2004)
4. Lin, J.-F., Drury, C., Karwan, M., Paquet, V.: A general model that accounts for Fitts' law and Drury's model. In: Proceedings of the 17th Congress of the International Ergonomics Association, Beijing, China, August 9-14 (2009)
5. Lin, J.-F., Drury, C.G.: Modeling Fitts' law, in Ergonomics for all: celebrating PPCOE's 20 years of Excellence. In: Lin, D.-Y.M., Chen, H.-C. (eds.) Selected Papers of the Pan-Pacific Conference on Ergonomics, pp. 561–567. CRC Press, Taiwan (2010)
6. Gan, K.-C., Hoffmann, E.R.: Geometrical conditions for ballistic and visually controlled movements. Ergonomics 31, 829–839 (1988)
7. Novak, K.E., Miller, L.E., Houk, J.C.: Kinematic properties of rapid hand movements in a knob turning task. Experimental Brain Research 132, 419–433 (2000)
8. Milner, T.E.: A model for the generation of movements requiring endpoint precision. Neuroscience 49(2), 487–496 (1992)
9. Howarth, C.I., Beggs, W.D.A., Bowden, J.M.: The relationship between speed and accuracy of movement aimed at a target. Acta Psychologica 35, 207–218 (1971)
10. Lin, R.F., Ho, Y.-C.: Verification of ballistic movement models in a true 3D environment. In: Proceedings of the 2nd East Asian Ergonomics Federation Symposium, October 5-8. National Tsing Hua University, Hsinchu (2011)
11. Lin, R.F., Shih, S.-W., Jiang, B.C.: An application of ballistic movement models for comparing ageing differences while interacting with a touchscreen. In: Proceedings of the 4th International Conference on Applied Human Factors and Ergonomics, San Francisco, California, USA, July 21-25 (2012)
12. MacKenzie, I.S.: A note on the information-theoretic basis for Fitts' law. Journal of Motor Behavior 21, 323–330 (1989)
13. MacKenzie, I.S.: Fitts' law as a research and design tool in human-computer interaction. Human-Computer Interaction 7(1), 91–139 (1992)

AugmentedBacklight: Expansion of LCD Backlights Using Lighting Methods in the Real World

Maho Oki[1], Koji Tsukada[1,2], and Itiro Siio[3]

[1] Future University Hakodate, Hokkaido, Japan
{okimaho,tsuka,siio}@acm.org
[2] JST PRESTO, Saitama, Japan
[3] Ochanomizu University, Tokyo, Japan

Abstract. People experience various lights in daily life, such as sunlight through leaves and candle light. However, digital contents shown in common LCDs cannot express these impressive lights. In this paper, we propose a novel display technique, AugmentedBacklight, which can express impressive lights in the real world by expanding the LCD backlight. The system mainly consists of a transparent LCD, a screen behind the LCD, and a projector behind the above devices. The projector shows videos of impressive lights on the screen for the backlight of the transparent LCD. Thus, users can see digital contents (e.g., e-books and pictures) under various impressive lights in the real world.

Keywords: Display, backlight, lighting.

1 Introduction

People experience various lights in daily life, such as sunlight through leaves, candle light, and sun glare on water. These lights affect appearances of objects in the real world. For example, people often feel different impressions by reading books under sunlight through leaves compared to under fluorescent lamps. However, digital contents shown in common LCDs cannot express these impressive lights since the LCD backlight emits flat lights –same brightness in each pixel – to keep visibility. To overcome this limitation, we propose a novel display technique, AugmentedBacklight, which can express impressive lights in the real world by expanding the LCD backlight.

2 Related Works

Seetzen et al. [5] and Chen et al. [1] proposed novel displays that have similar mechanisms with our system: that is, integrating a transparent LCD and a DLP projector/LED-based display to expand the LCD backlight. Suyama et al. [6] proposed a Depth-fused 3D (DFD) display that enables a user to perceive an apparent 3D image using overlaid two displays. Lee et al. [4] proposed a unique DFD display that integrates fog screen and two projectors.

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 209–216, 2013.

While these systems focused on expanding dynamic ranges of pictures/videos shown on the LCD or generating 3-D image, our system focuses on expressing appearances of contents under various impressive lights in the real world.

There were several art works [2, 3] that add visual effects to real time videos depending on activities of people in the real world. Our system may effectively work in such applications.

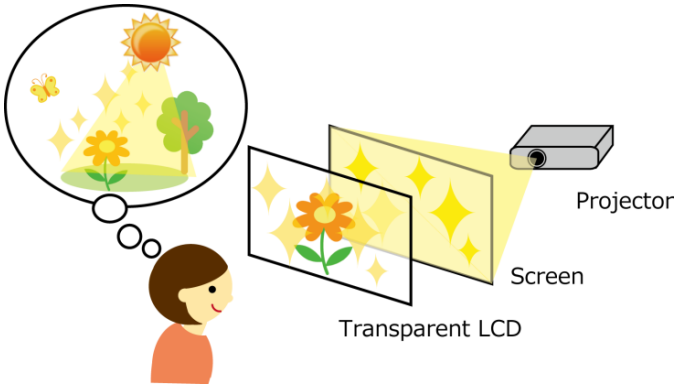


Fig. 1. Basic concept of AugmentedBacklight

3 AugmentedBacklight

The basic concept of AugmentedBacklight is to express various lights in the real world by expanding the LCD backlight (Fig. 1). Many lights in the real world (e.g., sunlight and spot lights) have so high luminance that most people have difficulty to look at them directly. These lights also have fluctuation depending on surrounding environment (e.g., sunlight reflected on the water). In this research, we aim to express these lights on the LCD using lighting video that includes impressive lights in the real world. For example, our system may recreate an experience of reading a book under sunlight through leaves by integrating texts and the video of such lighting.

The system mainly consists of a transparent LCD, a screen behind the LCD, and a projector behind the above devices. As shown in Fig. 1, the projector shows videos of impressive lights on the screen for the backlight of the transparent LCD. Thus, users can see digital contents (e.g., e-books and pictures) under various impressive lights in the real world.

3.1 Prototype

We developed a prototype as shown in Fig. 2. It consists of a projector (EPSON EB-410W, 2000lm), a transparent LCD (TCR-3104M) and a screen (handmade acrylic panel). The LCD and the projector are connected to a PC (Windows XP) via an USB-VGA convertor (Buffalo, GXDVI/ U2AI) and a built-in VGA connector.

The main program running on the PC controls (1) videos of impressive lights shown on the screen through the projector and (2) main contents (e.g., e- books, pictures or web pages) on the LCD. In this paper, we defined the prior contents as "light contents" and the later contents as "main contents". Since impressive lights in the real world (e.g., sunlight through leaves) often change their brightness and textures dynamically, we selected videos of such changing lights as the light contents.

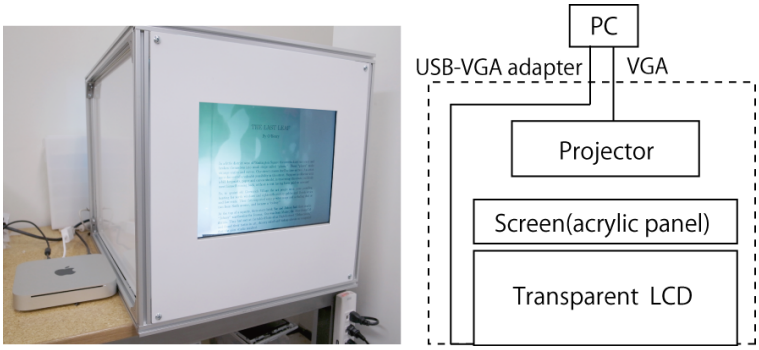


Fig. 2. The AugmentedBacklight prototype

3.2 Design of Light Contents

We selected light contents in consideration that users can perceive impressive lights on the display without bothered by uncomfortable feelings. In this paper, we focused on mainly two categories: (A) impressive lights in the real world, and (B) impressive visual effects created by computer graphics. The former category includes videos of sunlight through leaves thrown on the wall and sunlight reflected on water. The latter category includes visual effects of lightning and brilliant lights. These examples of light contents are shown in Fig. 3. These contents are converted into gray scale to avoid affecting colors of main contents.

3.3 Design of Main Contents

We selected still images as main contents to verify the basic performance of light contents. For example, we prepared pictures of "green leaves", "a flower", "lake" and "texts in an e-book" in corresponding to the impressive lights in the real world. We also prepared pictures of "people" and "a glass besides the window" suited for the visual effects. These examples of main contents are shown in Fig. 4.

4 Preliminary Study

We performed a preliminary study to verify the basic performance of the AugmentedBacklight prototype. We have selected 5 subjects (3 females, aged between 23 and 33) from our laboratory.

4.1 Procedure

We explain basic procedure of the preliminary evaluation. First, we set each subject down in front of the AugmentedBacklight prototype and explained the basic function of the system. Next, the experimenter shows a main content along with the white background instead of a light content. Then, the light contents are changed several times. After the subject repeated these processes for all main contents, the experimenter ask them to write the impression of the system.

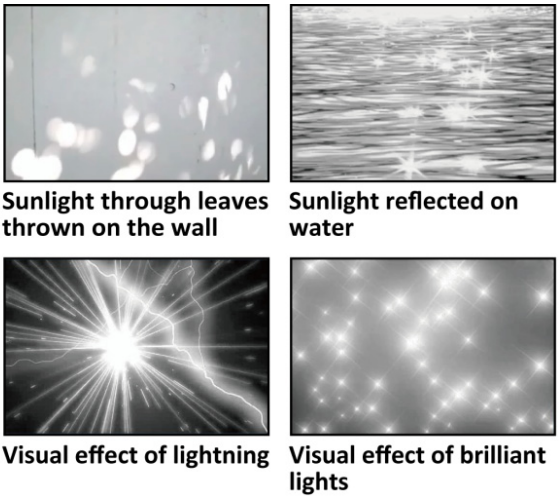


Fig. 3. Examples of light contents

Fig. 4. Examples of main contents

4.2 Results

In this section we explain the results from the viewpoints of “effective examples” and “findings”.

Effective Examples

First, we describe effective combinations of light contents and main contents. The picture of each combination is taken with a digital camera (GR Digital 2) at a close range.

Fig. 5 shows the combination of "green leaves" and "sunlight through leaves". The typical impressions written by the subjects were as follows:

- The green leaves look vivid by the movement of sunlight through leaves.
- Beautiful. The backlight looks actual sunlight through leaves.

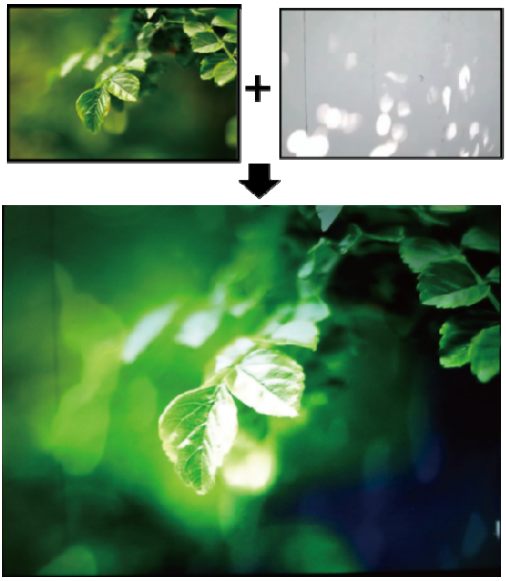


Fig. 5. The combination of "green leaves" and "sunlight through leaves"

Fig. 6 shows the combination of "texts in the e-book" and "sunlight through leaves". The typical impressions written by the subjects were as follows:

- The gentle movement of the sunlight brought me comfort.
- I can read texts more easily than I imagined.

Fig. 7 shows the combination of "surprised person" and "lightning effect". The impressions of the subjects were as follows:

- Strange feeling. The person and other objects look floating on the air.
- Cool. The person seems to emit the aura.
- I felt retro impression like old films or stained glasses.

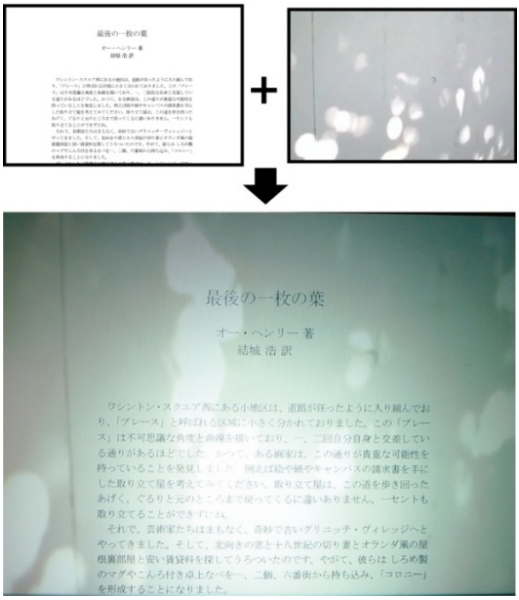


Fig. 6. The combination of "texts in e-book" and "sunlight through leaves"

Fig. 7. The combination of "surprised person" and "lightning effect"

Fig. 8 shows the combination of "a glass besides a window" and "brilliant lights". The impressions of the subjects were as follows:

- The picture looks like a 3D image. It seems to pop out from the display.
- I was interested in the effect since I have never seen such lighting in the real world.

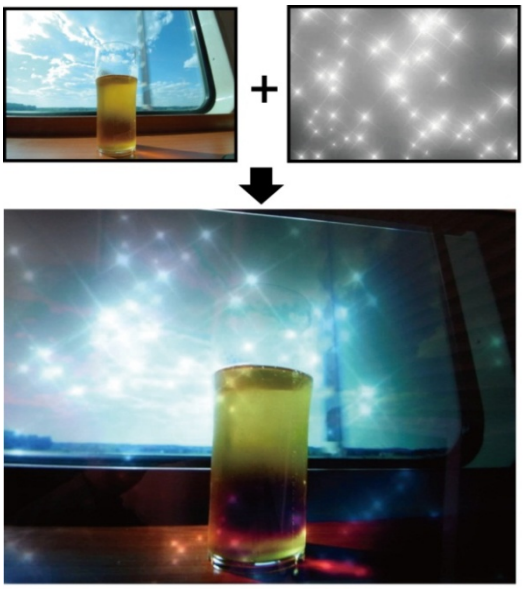


Fig. 8. The combination of "a glass besides a window" and "brilliant lights"

Findings

All subjects reported that the pictures shown with the light contents look different from ones with the white background. Some subjects felt the depth between light contents and main contents. Others reported each main contents looks drastically different depending on light contents.

Since most subjects felt the stereo effect in the picture of "a glass besides a window" in which the main object is simply located on the center, the light contents may effectively works with pictures that have simple compositions.

As the combination of light/main contents, most subjects had comfort with the combination of contents in similar situations. For example, "green leaves" and "sunlight through leaves" as shown in Fig. 5 are one of the most popular examples in the preliminary evaluation. On the other hand, some light contents (e.g., "sunlight through leaves") work rather effectively with variety of main contents. For example, a subject mentioned "I would like to use the light content (sunlight through leaves) as the background of my desktop".

Meanwhile, visual effects (e.g., lightening and brilliant lights) were not suited for simple main contents (e.g., texts in e-book) since these effects have too strong presence that most subjects have difficulty to read texts with them. However, when the main contents are carefully selected, the visual effects may give great impacts on users as shown in Fig. 7.

In this study, we confirmed some effective examples of contents for the AugmentedBacklight. We would explore people's impressions on various combinations of light/main contents more accurately. Moreover, we also try to specify the cognitive factor that affects unique experiences with our system.

5 Conclusion

In this paper, we proposed a novel display technique, "AugmentedBacklight", which can express impressive lights in the real world by expanding the LCD backlight. We have developed a prototype system using a transparent LCD, a screen behind the LCD, and a projector behind the above devices. We introduced the examples of light contents shown on the screen through the projector and main contents shown on the LCD. We also explained the effective examples through the preliminary evaluation.

Acknowledgements. This work was supported by JST PRESTO program.

References

1. Chen, H., Sung, J., Ha, T., Park, Y., Hong, C.: Backlight local dimming algorithm for high contrast LCD-TV. In: Proc. of ASID 2006, pp. 168–171 (2006)
2. Fukuchi, K., Mertens, S., Tannenbaum, E.: EffecTV: a real-time software video effect processor for entertainment. In: Entertainment Computing–ICEC 2004, pp. 341–406 (2004)
3. Iwai, T.: Piano–as image media. Leonardo 34(3), 183–183 (2001)
4. Lee, C., DiVerdi, S., Höllerer, T.: An immaterial depth-fused 3D display. In: Proceedings of the 2007 ACM Symposium on Virtual Reality Software and Technology, pp. 191–198 (2007)
5. Seetzen, H., Heidrich, W., Stuerzlinger, W., Ward, G., et al.: High dynamic range display systems. ACM Transactions on Graphics (SIGGRAPH 2004) 23(3), 760–768 (2004)
6. Suyama, S., Ohtsuka, S., Takada, H., Uehira, K., Sakai, S.: Apparent 3-D image perceived from luminance-modulated two 2-D images displayed at different depths. Vision Research 44(8), 785–793 (2004)

An Exploratory Study to Understand Knowledge-Sharing in Data-Intensive Science

Jongsoon Park and Joseph L. Gabbard

Virginia Bioinformatics Institute, Industrial & Systems Engineering
Virginia Tech
{babirong,jgabbard}@vt.edu

Abstract. This paper describes a challenges associated with data-intensive research processes, knowledge-sharing phenomena, and end-users' expectations in the field of bioinformatics. We developed a questionnaire to support deeper understanding of user experiences with knowledge sharing activities. The results reveal that there are several challenging issues biologists encounter when using bioinformatics resources. A much smaller number of biologists have engaged in passive knowledge sharing within their research fields than we had expected. However, most biologists expressed their willingness to share their own knowledge with others. This result reinforces the need for more user-centered design approaches for supporting knowledge-sharing in rapidly emerging fields of data-intensive science. At the same time, our results suggest that more work is needed to examine how to best motivate users to further engage and contribute knowledge in online scientific communities.

Keywords: User-centered Design Approach, Knowledge-Sharing, Data-Intensive Science.

1 Introduction

In the past two decades, we have seen an exponential increase in the size and breadth of available scientific data, demanding new integrated solutions to explore and elicit valuable insights more efficiently (Kelling et al., 2009). As such, there is growing interest in data-intensive scientific discovery through data integration, simulation, visualization, and validation across distributed networks of heterogeneous resources. A notable example is in biological sciences, which is currently undergoing a rapid paradigm shift to data intensive science (Bell, Hey, & Szalay, 2009). Diverse bioinformatics resources (e.g. online resources that integrate biological data and analysis tools) have been developed, potentially enabling biologists to analyze huge and complex data sets faster and more efficiently as compared to traditional methods (Bull, Ward, & Goodfellow, 2000; Katoh, 2002; Yarfitz, 2000). However, most of these resources have been developed by applied scientists (i.e. computer scientists and bioinformaticians), and are not intuitive or rich enough to address most bench biologists' goals. At the same time, bench biologists are beginning to appreciate the power and potential of bioinformatics resources, despite their poor usability.

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 217–226, 2013.

Generally speaking, biologists and other researchers in data-intensive fields are grappling with how best to deal with "big data" while HCI researchers (collaborating with bioinformaticians) are working to understand how best represent and interact with such data. The question remains: how can we best close this gap to ensure that these resources are both powerful and intuitive to varying user classes. There have been various attempts to bridge gaps between technology-driven bioinformatics resources and the broader spectrum of biologists' work practices, such as creating more goal-oriented solutions for data collection and storage, and conducting task analysis and usability studies (Joan Bartlett, Ishimura, & Kloda, 2011; J. Bartlett & Neugebauer, 2005; D. Bolchini, 2009; Davide Bolchini, 2009; Javahery, 2004; Mirel, 2009; Tran, Dubay, Gorman, & Hersh, 2004). In parallel, we see web-based "knowledge-sharing platforms" as a growing trend to support data-intensive discovery research by allowing scientists to exchange data, ideas, expertise, and scientific literature online to improve the effectiveness of their processes and validity of their outcomes (De Roure, Goble, & Stevens, 2009; Li, 2012; McIntosh et al., 2012; Parnell, 2011). These knowledge-sharing platforms in scientific communities may provide a timely mean to assist biologists working with large and consistently growing diverse data sets. However, to date, little attention has been paid to a comprehensive understanding of end-users' characteristics and attitudes about knowledge sharing based on the culture of these data-intensive research domains.

To support useful and seamless knowledge-sharing and reuse in data-intensive research, we need to examine a series of higher level questions from the user's perspective, such as: What are the major shortcomings in current online bioinformatics resources? To what extent do end-users have experience with knowledge-sharing activities (e.g., knowledge-sharing and -reuse)? What are end-users' expectations of knowledge-sharing activities? What are users' perceptions of challenges and opportunities in knowledge-sharing environments? Examining these questions will provide us with meaningful insights not only to identify unmet needs and opportunities, but also to support cross-disciplinary scientific research in data-intensive fields such as biology.

2 Objective and Research Question

This study has three objectives: 1) to understand end-users' perspectives on shortcomings of current online bioinformatics resources, 2) to identify to what extent end-users have experience with knowledge-sharing activities to support their research processes, and, 3) to elicit specific end-users' concerns and expectations of knowledge-sharing.

3 Methods

We developed a questionnaire to elicit users' experience with, perceptions of, and attitudes towards knowledge-sharing activities. The questionnaire is based on previous studies (Bock, Zmud, Kim, & Lee, 2005; Preece, Nonnecke, & Andrews,

2004) and feedback from domain experts who are familiar with biological "wet-lab" experiments and have worked in fields of biology for five years or more.

The first set of questions elicits participant demographics (e.g., age, gender, current work/academic role) and background (e.g., usage frequency of bioinformatics resources). The second set of questions elicits information on various aspects of users' experiences, challenges, and expectations with current bioinformatics resources and knowledge-sharing related to their research processes.

After data were collected, we employed statistical analysis to describe characteristics and behavior of users' current knowledge-sharing activities. The findings of this survey are intended to help identify and prioritize distinguishing web resource features needed to support online knowledge-sharing in data-rich scientific processes.

3.1 Participants

We collected responses from participants of workshops offered by the Virginia Bioinformatics Institute and from PhD students in the Virginia Tech Genomics, Bioinformatics, Computational Biology graduate program. However, the calculation of a response rate was difficult because we do not know how many total PhD students in the program were invited to participate. Eighty-one of eighty-four total responses were usable (three were incomplete or incomprehensible).

More than half of the participants (55.6%) are male, and 72.9% are between 20 and 39 years of age. Almost all characterized their main research role as biologist (63), with others self-reporting roles of bioinformatician (6), chemist (4), computer scientist (4), mathematician (3), and other (11) such as microbiologists, biochemists, and clinician (note: participants were asked to select all items that apply). Slightly over 60% of participants (n = 49) have over five years of research experience in biology, while about 40% (n = 32) have been conducting research for no more than five years.

The frequency of use of bioinformatics resources ranges from every day to less than once a month. A majority of participants (70.3%) reported using bioinformatics resources more than once a week, 13.6% reported use as more than once a month, and 13.6% reported using bioinformatics resources around once a month.

3.2 Results

We performed data analysis using SPSS (version 18.0), defining statistical significance at $\rho < 0.05$. In general, we observed similar response patterns among participants with no more than 5 years of research experience. Those with over 5 years research experience also showed similar response tendencies. As such, in the following discussion, we consider two broad classes of participants; those with no more than 5 years of research experience and those with over 5 years of research experience.

3.3 Limitations in Online Bioinformatics Resources

Our results show that most participants are currently challenged by the lack of integration and inconsistent results across online bioinformatics resources (e.g. different gene naming conventions, different annotations for the same gene). In the same vein, they repeatedly highlighted limitations due to the poor quality of genomic sequences and metadata. Some participants noted strengths of bioinformatics resources such as multiple views on the same data and multiple comparisons across different genomes. Inconsistency in user interfaces and general lack of usability were cited as major difficulties for a number of participants, implying a steep learning curve (i.e., long learning times) as a key usability issue. In addition, some participants had trouble accessing data due to complex navigation structures typical of bioinformatics resources. Lastly, data security was noted as an important issue, since many researchers are leveraging these resources to support hypotheses generation, publications or grants.

3.4 Important Factors of Bioinformatics Resources

To examine users' expectations of bioinformatics resources, we asked participants which resource features are the most important or valuable. Multiple responses were categorized and tallied using the multiple-dichotomy frequency analysis. We constructed a cross tabulation table to analyze the most dominant participants' response. Table 1 presents the cross-tabulation frequencies by years of research experience.

Table 1. Important factors of bioinformatics resources

	Research Experience (yrs)		**Total**
	No more than 5	**over 5**	
Speed and responsiveness of resource	23 (71.9%)	36 (76.6%)	59 (74.7%)
Wealth of available data	22 (68.8%)	30 (63.8%)	52 (65.8%)
Breath of resource tools and functions	16 (50.0%)	29 (61.7%)	45 (57.0%)
Degree of data integration	17 (53.1%)	26 (55.3%)	43 (54.5%)
Ease of use	16 (50.0%)	26 (55.3%)	42 (53.2%)
Ability to upload my own data	13 (40.6%)	25 (53.2%)	38 (48.1%)
Ability to ask questions related to my research	12 (37.5%)	23 (48.9%)	35 (44.3%)
Ability to create publication quality images	10 (31.3%)	22 (46.8%)	32 (40.5%)
Advanced visualizations	13 (40.6%)	19 (40.4%)	32 (40.5%)
Ability to collect knowledge from others researchers	9 (28.1%)	21 (44.7%)	30 (38.0%)
Ability to share knowledge with other researchers	6 (18.8%)	9 (19.1%)	15 (18.9%)
Total # of participants	32	47	79

As expected, performance-related factors common to most web-based systems ranked relatively high in "important factors of bioinformatics web resources". Namely, participants valued "high speed and responsiveness of resource" (74.3%), followed by "wealth of available data" (65.8%), "breadth of resource tools and

functions" (57.0%), "degree of data integration" (54.5%), "ease of use" (53.2%), and "ability to upload my own data" (48.1%).

Interestingly, nearly half of the participants with over 5 years of research experience selected "ability to ask questions related to my research" (48.9%) and "ability to collect knowledge from others" (46.8%) as an important resource features. In contrast, a much smaller proportion of these participants (19.1%) appear interested in sharing their knowledge with others. It can be inferred from these results that experienced participants are more interested in enhancing the overall quality (and performance) of their research by making use of others' shared knowledge than sharing their accumulated knowledge and skills. In comparison with the above findings, participants with no more than 5 years of research experience showed little interest in sharing and collecting knowledge as compared to other features.

3.5 Knowledge Sharing Experience

Our knowledge-sharing results suggest significant, but limited, online knowledge-sharing activity among our sampled user population (See Fig. 1).

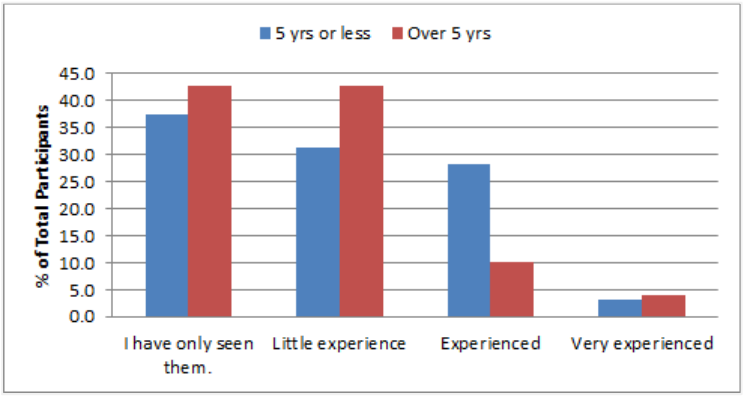


Fig. 1. Knowledge sharing experience

Specially, all participants reportedly engage in knowledge-sharing activities to some degree (e.g., knowledge-sharing or -reuse), but nearly 80 percent of respondents reported participating in knowledge-sharing activities in a passive manner. That is, most users rarely share their knowledge, but instead tend to seek and use knowledge shared by others. This is a typical example of a lurker; an individual who consumes information and requests specific questions from others, but does not explicitly contribute to the shared knowledge base (Preece et al., 2004).

To examine potential correlations among age, years of research experience, and experience level in knowledge-sharing activities, we performed a Kendall's tau correlation analysis. As might be expected, we found a positive, statistically significant correlation between age and research experience (Kendall's tau $b = 0.608$, $\rho = 0.000$). Contrary to our expectation, however, we found no significant correlation

between years of research experience and the experience level in knowledge-sharing activities (Kendall's tau b = -0.154, ρ = 0.123). Moreover, there was no correlation between age and knowledge-sharing experience (Kendall's tau b = -0.019; ρ = 0.847). Even though one might expect younger researchers to be more attuned to online knowledge-sharing opportunities.

We next examined the types of knowledge consumed in knowledge-sharing activities; including specific examples of both implicit and explicit knowledge (Bock et al., 2005; Choo, 2000). These results provide valuable insight into what types of knowledge participants seek from others in support of their research (See Table 2).

Table 2. Types of shared knowledge employed

		Research Experience (yrs)		Total
		no more than 5	over 5	
	None	0 (0%)	6 (18.8%)	6 (11.1%)
Implicit Knowledge	General ideas	14 (63.6%)	15 (46.9%)	29 (53.7%)
	Accumulative research experiences	9 (40.9%)	15 (46.9%)	24 (44.4%)
	Unique opinions	6 (27.3%)	7 (21.9%)	13 (24.1%)
Explicit Knowledge	Articles published in books, websites, and documents	13 (59.1%)	16 (50.0%)	29 (53.7%)
	Products, patents, databases, tools, and prototypes	8 (36.4%)	7 (21.9%)	15 (27.8%)
	Rules, routines, or operating procedures	10 (45.5)	5 (15.6%)	15 (27.8%)
	Total # of participants	22	32	54

The most frequently reported knowledge used is “general ideas” (53.7%), “articles published in books, websites, and documents” (53.7%), followed by “accumulative research experiences” (44.4%). The overall pattern of responses indicates that our participants employ implicit and explicit knowledge in an evenly balanced way.

We also found that participants with no more than 5 years research experience tended to rely on "general ideas" (63.6%) more than participants with over 5 years research experience (46.9%), with open-ended responses suggesting that the relatively inexperienced participants use shared knowledge extensively to generate research questions and confirm hypotheses. In addition, these relatively inexperienced researchers depend on others to confirm rules or procedures (45.5%) as compared with more experienced participants (15.6%).

3.6 Knowledge Sharing Intention

Next, we assessed participants’ intention to share knowledge using questions adapted from Bock et al. (2005). Participants indicated their agreement or disagreement with statements using a seven-point Likert-type scale (where scores of 7 suggest a strong willingness to share). Surprisingly, almost all participants (95%) reported a willingness to share knowledge with others (Fig. 2). This result suggests there are

ample opportunities to promote and grow knowledge-sharing in data-intensive sciences such as biology.

Fig. 2. We used a one-way MANOVA to determine whether there are any differences between our two research experience groups on more than one kind of knowledge. We found no significant group effects for the types of knowledge on their intention to share knowledge, F (14, 117) = 0.738, ρ =0.743; *Wilk's* λ = 0.728, *partial* $\varepsilon 2$ = 0.54.

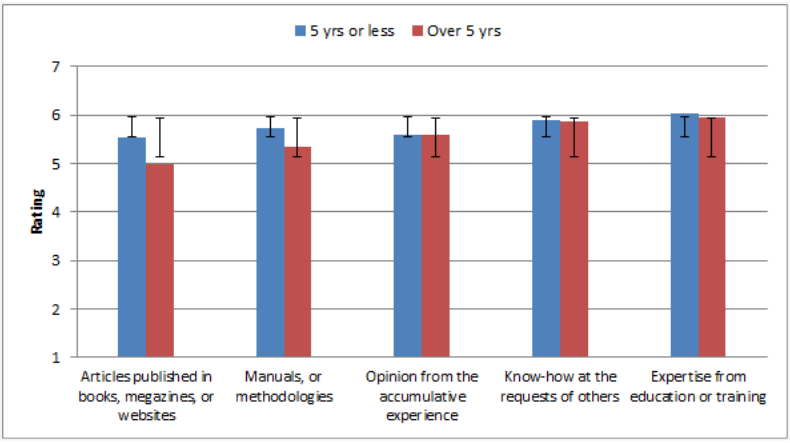


Fig. 2. Intention to share specific knowledge (7=strongly agree, 1=strongly disagree)

Next, we used an independent t-test to investigate differences in knowledge-sharing intentions between the two research experience groups, and found no significant differences between them across any of the five variables: articles published in books, magazines, or website, t(75) = 1.539, ρ = 0.128; manuals, or methodologies, t(73) = 1.316, ρ = 0.192; own opinion from the accumulative experience, t(75) = -0.023, ρ = 0.982; know-how at the requests of others, t(75) = 0.124, ρ = 0.901; and expertise from education or training, t(75) = 0.314, ρ = 0.755.

3.7 Concerns about Knowledge Sharing

One of the significant results to emerge from our responses is that most participants worry about source reliability (i.e. poster's expertise, data integrity, experimental verification, quality of information). Similarly, results indicate concerns about encountering "incorrect information or annotations". Other concerns include the need for an easy way to contribute their own knowledge, suggesting that participants have encountered usability issues when they have previously tried to share information. Additionally, participants noted response time, frequency of knowledge update, copyright, and privacy as potential barriers to engaging in knowledge-sharing activities.

3.8 Expectations

We found the following three themes of expectations concerning knowledge-sharing from the free text comments. First, there were many responses that referred to source creditability. For example, participants expect to be able to authenticate contributors' expertise as well as shared knowledge. Second, about 50% of participants said they often need technical support to make use of bioinformatics tools (e.g. troubleshooting advice, application tips, how-to guides). Participants noted the lack of well-organized information repositories of shared knowledge by domain, technique, methodology, etc. Third, there was some evidence that participants are interested in reusing knowledge from others to improve their research processes and outcomes. Frequently mentioned expectations include implicit knowledge related to protocols used for conducting biological experiments such as, small "bench work tricks", protocols of best practices for cutting edge -omics research, and information about negative data.

4 Discussion

The results of this research indicate not only existing barriers to use of bioinformatics resources but also opportunities to address users' unmet needs in data-intensive scientific communities. We found no significant differences across age groups and years of research experience; suggesting that designs to support knowledge sharing should consider other user class characteristics.

Our results indicate that biologists struggle to utilize bioinformatics resources regardless of years of research experience mainly due to inconsistent results and poor user interfaces. Moreover, users indicate that the greatest shortcomings of current resources are often associated with the most important features. A possible explanation may be that many online bioinformatics resources employ a system-oriented development approach rather than user-centered design approach that aims to better understand users' unmet needs.

We found approximately 80% of participants do not actively engage in knowledge-sharing than we expected, regardless of age or years of research experience. Only a small portion of participants have actively engaged in research-related knowledge-sharing. A possible explanation might be that current links exist between knowledge-sharing platforms and biologists are limited or ad hoc despite the prevalence of online knowledge-sharing resources in scientific communities. In other words, we can suppose that many knowledge-sharing platforms are built with a focus on current technological trends rather than user experience factors, which affect users' motivation to engage in knowledge-sharing activities. Another possible explanation is that characteristics of the scientific culture may influence an individual's propensity to engage in knowledge sharing and reuse. Thus, more study is needed to understand how best to foster knowledge sharing and reuse in scientific communities.

Results from this study also show that participants with no more than 5 years research experience rely more on implicit knowledge shared by other practitioners than explicit knowledge. It seems possible that self-efficacy caused by accumulated

expertise may contribute to the differences found in usage patterns of implicit knowledge.

The most significant finding is that nearly 95% of participants are willing to share their knowledge, contradicting the very low levels of current involvement in knowledge-sharing activities. The lack of quality user experiences to support seamless and easy knowledge contribution suggests a need to develop knowledge-sharing platforms that embrace user-centered design approaches.

Lastly, this work identifies several issues that must be ensured to facilitate knowledge-sharing and reuse in data intensive settings aiming to support scientific discovery. A majority of participants were concerned about the quality of knowledge and the degree to which they can trust shared knowledge. These results are consistent with those of previous studies (Golbeck, 2008; Levin, Cross, Abrams, & Lesser, 2002) and suggests that source credibility has a considerable impact on attitudes towards knowledge-sharing. More study is therefore needed to better understand how to cultivate trust in, and increase motivation to use, knowledge-sharing activities.

5 Conclusion

This research is one of the first studies to investigate knowledge-sharing in emerging data-intensive sciences such as biology. Our results imply significant opportunities to support knowledge-sharing in these communities, but that careful attention needs to be taken to users' perceived and actual needs. This initial study focuses on eliciting basic user experiences with, and perceptions of, bioinformatics resources and online knowledge-sharing activities. Results presented herein may inform future studies to explore user experiences and knowledge-sharing activities in data-rich environments.

This study was mainly conducted among biologists. The outcomes of the study could be strongly influenced by the culture of experimental science. Hence, to determine whether these findings can be applied to a wide range of knowledge-sharing platforms for rapidly emerging fields of data-intensive science, there needs to be further study with additional participants from different background (e.g. applied scientists) across other data-intensive fields (e.g., visual analytics, meteorology).

Acknowledgments. This work was supported in part with Federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services, under Contract No. HHSN272200900040C.

References

1. Bartlett, J., Ishimura, Y., Kloda, L.: Why choose this one?: Factors in scientists' selection of bioinformatics tools. Information Research 16(1), 15 (2011)
2. Bartlett, J., Neugebauer, T.: Supporting information tasks with user-centred system design: The development of an interface supporting bioinformatics analysis. Canadian Journal of Information and Library Science 29(4), 486–487 (2005)

3. Bell, G., Hey, T., Szalay, A.: Beyond the data deluge. Science 323(5919), 1297–1298 (2009)
4. Bock, G.W., Zmud, R.W., Kim, Y.G., Lee, J.N.: Behavioral intention formation in knowledge sharing: Examining the roles of extrinsic motivators, social-psychological forces, and organizational climate. MIS Quarterly, 87–111 (2005)
5. Bolchini, D.: Better bioinformatics through usability analysis. Bioinformatics (Oxford, England) 25(3), 406–412 (2009), doi:10.1093/bioinformatics/btn633
6. Bolchini, D., Finkestein, A., Paolini, P.: Designing Usable Bio-information Architectures. In: Jacko, J.A. (ed.) HCII 2009, Part IV. LNCS, vol. 5613, pp. 653–662. Springer, Heidelberg (2009)
7. Bull, A.T., Ward, A.C., Goodfellow, M.: Search and discovery strategies for biotechnology: the paradigm shift. Microbiology and Molecular Biology Reviews 64(3), 573–606 (2000)
8. Choo, C.W.: Working with knowledge: how information professionals help organisations manage what they know. Library Management 21(8), 395–403 (2000)
9. De Roure, D., Goble, C., Stevens, R.: The design and realisation of the Virtual Research Environment for social sharing of workflows. Future Generation Computer Systems 25(5), 561–567 (2009), doi:10.1016/j.future.2008.06.010
10. Golbeck, J.: Weaving a web of trust. Science 321(5896), 1640–1641 (2008)
11. Javahery, H.: Beyond power making bioinformatics tools user-centered. Communications of the ACM 47(11), 58 (2004), doi:10.1145/1029496.1029527
12. Katoh, M.: Paradigm shift in gene-finding method: From bench-top approach to desk-top approach (review). Int. J. Mol. Med. 10(6), 677–682 (2002)
13. Kelling, S., Hochachka, W.M., Fink, D., Riedewald, M., Caruana, R., Ballard, G., Hooker, G.: Data-intensive science: a new paradigm for biodiversity studies. t BioScience 59(7), 613–620 (2009)
14. Levin, D.Z., Cross, R., Abrams, L.C., Lesser, E.L.: Trust and knowledge sharing: A critical combination. IBM Institute for Knowledge-Based Organizations, 1–9 (2002)
15. Li, J.W.: SEQanswers: an open access community for collaboratively decoding genomes. Bioinformatics (Oxford, England) 28(9), 1272–1273 (2012)
16. McIntosh, B.K., Renfro, D.P., Knapp, G.S., Lairikyengbam, C.R., Liles, N.M., Niu, L., Hu, J.C.: EcoliWiki: a wiki-based community resource for Escherichia coli. Nucleic Acids Research 40(D1), D1270–D1277 (2012), doi:10.1093/nar/gkr880
17. Mirel, B.: Supporting cognition in systems biology analysis: findings on users' processes and design implications. BioMed Central Ltd. (2009)
18. Parnell, L.D.: BioStar: An Online Question & Answer Resource for the Bioinformatics Community. PLoS Computational Biology 7(10), e1002216 (2011)
19. Preece, J., Nonnecke, B., Andrews, D.: The top five reasons for lurking: improving community experiences for everyone. Computers in Human Behavior 20(2), 201–223 (2004), http://dx.doi.org/10.1016/j.chb.2003.10.015
20. Tran, D., Dubay, C., Gorman, P., Hersh, W.: Applying task analysis to describe and facilitate bioinformatics tasks. Research Support, N I H, Extramural (2004)
21. Research Support, U S Gov't, P H S. Stud Health Technol. Inform 107(Pt 2), 818–822
22. Yarfitz, S.: A library-based bioinformatics services program. Bulletin of the Medical Library Association 88(1), 36 (2000)

Interaction of the Elderly Viewer with Additional Multimedia Content to Support the Appreciation of Television Programs

Kamila Rodrigues, Vânia Paula de Almeida Neris, and Cesar Augusto Camillo Teixeira

Department of Computer Science, Federal University of São Carlos/UFSCar,
São Carlos, Brazil
{kamila_rodrigues,vania,cesar}@dc.ufscar.br

Abstract. For many people, television is still the main form of entertainment. The elderly population, in particular, spends much time at home and uses the television for companionship and entertainment. However, in some countries, due to low education level or the absence of it, to physical limitations and/or cognitive difficulties, to the lack of prior knowledge, among other difficulties, part of the elderly viewers do not enjoy enough the programs to have satisfaction with them. With the interactivity provided by the Digital TV, additional information may help the elderly viewer to better appreciate television programs and have more fun. This paper presents data from an observational case study conducted in the Brazilian scenario to analyze the interaction of elderly viewers with TV programs enriched with additional multimedia content in different formats. From the experience, some good practices for the design of additional multimedia content to the elderly viewer were formalized.

Keywords: Design, Interactive Digital TV, Elderly, Additional Multimedia Content, Satisfaction.

1 Introduction

Technological advances and facilities for the dissemination and exchange of information have prompted people to a more engaging socialization. However, part of society faces obstacles to fully enjoy the benefits offered by Information and Communication Technologies (ICTs). The reasons for this include economical, educational, cultural and physical or cognitive problems. Part of the elderly population also presents difficulties on handling of new technologies, including ICTs.

In many countries, populations are aging. According to the United Nations (UN), in its latest technical report "*World Population Prospects*", in the next 43 years the number of people over 60 years old will be three times greater than the current one. The elderly will represent a quarter of the estimated world population, i.e., about 2 billion people (in a total of 9.2 billion) [1].

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 227–236, 2013.

However, some researches indicate that in fact, elderly people have found barriers to use ICTs [2]. Among the identified barriers are: the inability to read, due to the low level of education of some populations, and psychological fears because of the cost of some equipment [2]. The elderly user believes to be difficult to interact with new technologies and fears to make mistakes which could damage the equipment [2]. However, the same technology that inhibits may be used to get the best of the benefits it provides. The integration of efforts in different areas, including computing, can promote a better integration of the elderly in the society.

In Brazil, no other segment of the population watches TV as much as the elderly [3], [4]. With the interactivity provided by the interactive Digital TV (iDTV) and applying the concept of Universal Usability and Integral Accessibility [5], it is possible to provide specific services designed to the needs of specific population, as elderly viewers, for example. These services, which generally require extra efforts and investments, may not result in additional costs for TV broadcasters and viewers if implemented appropriate business models.

Digital TV allows independent manipulation of different media that can be properly synchronized for a multimedia presentation. This feature facilitates the incorporation of content into television programs. The added content may be referenced and non-intrusive. The possibility of representing the additional content in different media is a key feature to meet different types of viewers' needs and their particularities. In this paper, the term "*additional multimedia content*" refers to extra information (audio, image, text, etc.) added to the television media in order to promote something extra to the viewer, aiming to clarify, inform or suggest, synchronously with the television program in question.

Offering additional multimedia content to television programs is also a way to provide accessibility to the elderly public, thus contributing to their involvement with new technologies and promoting greater inclusion into the society. Such technological solution based on additional information may help the elderly viewer to better appreciate television programs and, hopefully, have more fun.

This paper presents data from an observational case study conducted to analyze the interaction of elderly viewers with TV programs enriched with additional audio, text and animation. Elderly of a community center were invited to watch a segment of a documentary program and twelve of them volunteered to participate in the case study. The main objective was to analyze the interaction of the elderly viewer with additional multimedia content offered to television programs and to identify which additional content is best suited for this audience. With the experience, some good practices for the design of additional content to the elderly viewer were formalized.

This paper is divided as follows: section 2 reports surveys that provide information about the interaction of elderly people with television programs and the difficulties that they present due to the age; section 3 depicts some aspects of Digital TV, beyond to emphasize the inclusion of additional content in television programs as a way to provide additional information to the main content; section 4 describes the case study conducted with the elderly audience; section 5 discusses the main obtained results with the case study; section 6 presents some lessons learned and good practices for the design of interactive additional content for the elderly viewers; the conclusions are presented in section 7.

2 Television and the Elderly

Television plays several roles in the elderly lives. It is probably the main source to keep them informed of news, besides being a way to be entertained, to look for subjects to chat and share with friends, to relax, to relieve their loneliness, to prevent boredom, among others [6].

Surveys show that news programs, comedy, drama, soap operas, sports and music, are the preferred programs among elderly viewers [6], [7]. However, the pleasure and satisfaction provided by TV can be reduced if the viewers acquire physical and cognitive difficulties, typical of elderly people, such as the loss of hearing, vision or understanding of the content broadcasted.

Aging process promotes reductions in the efficiency of the sense organs [8]. From the cognitive point of view, some researches indicate that it is common that memory failures occur with advancing age and it is generally easier for the elderly to remember information when memory is tested by the recognition than when it is tested by recall [8].

Carmichael [8] reports in the guide for the development of interactive TV services for elderly viewers that, to design an appropriate solution for the elderly, it is necessary to know this public and the difficulties acquired due to age. Important issues should be considered when designing interactive television services for these viewers. They are: a) Elderly suffer visual and auditory changes as they age and these change decreases the accuracy of visual reception and also of auditory information obtained; b) Speech should be presented clearly and accurately, to avoid a negative impact on elderly viewers; c) Elderly have different changes during aging and, as user, demand different requirements therefore, they cannot be considered a homogeneous public. It should be careful with solutions that require dexterity and fine motor coordination, such as those based on touch screens; d) Some important information can be lost, because the nervous system of the elderly tends to process data received by the sensor organs more slowly; e) Elderly forget acquired information more quickly. Situations that require memorization should be replaced by the presentation of situations familiar to the elderly; and, f) The fear of the elderly to use new technologies can be minimized or eliminated if they understand the personal benefits they can get. It is also important that the first experience in using the technology be positive and successful.

In addition to potential commercial interests, the social responsibility requires proposals of new ways to design interactive multimedia systems to meet the diversity of the elderly public. From this perspective, it is a challenge to exploit the flexibility offered by the new TV, digital and interactive, to provide accessibility services, using different broadcasted media independently and simultaneously to the main program.

3 Interactive Digital TV and Additional Multimedia Content

Digital TV systems provide higher definition imaging and video, interactivity, and new services [9]. Additional media content (audio, video etc), and applications that

synchronize, coordinate the presentation of various media and manage the interactions promoted by the viewers can be sent with the main media (TV program). Examples of these applications are Electronic Program Guide (EPGs), banking, educational services, etc. Digital TV also enables the creation of nonlinear television programs [10].

It is possible to exploit the flexibility offered by the digital TV, that allows other media to be presented synchronized to the main program, and provide additional content to the viewer with more information and explanations about the program. This solution may offer, at the end of the program, more satisfaction to the viewers [11].

Current researches about the use of additional content as a form to enrich television programs extend the paradigm of TV. These researches are based on the concept of Social TV which explores the approach of consumption and sharing of television programs with edition and reuse of the content by viewers [12]. In this kind of research viewers may enrich parts of programs with information or indications and share them with other groups or communities.

There are also researches that combine interactive Digital TV with hyper video. In these cases the additional information are associated with the objects presented in scenes. Hyperlinks are inserted into the video and users access the information selecting a point on a single image area [13], [14].

Although these solutions are interesting, none of them considers the support for different types of viewers and their particularities. Also, they do not consider the difficulties and rejection of the elderly viewer, for example, when using new technologies. Therefore, it is necessary to define proper contents to be added to the TV program and also to define the best way to present them. The additional content must meet the needs of different types of viewers and to follow the patterns of production and exhibition of TV programs [15].

In the context of this work, it is necessary to think about a design process that considers the elicitation, development, presentation and validation of additional content, which is flexible and allows the offer of additional content for different types of viewers, including the elderly. However, to provide services centered on users' needs and with the application of the concepts of Universal Usability and Integral Accessibility, it is necessary to get closer to the users and identify features that meet their interaction requirements. It is also necessary to identify how to present, in the user interface, these features preventing the elderly to reject the solution. The interactive experience can bring more satisfaction if these issues are considered in the design of process of television content.

In this scenario, this paper presents a case study conducted with elderly users. This study observed the relationship of those users with an enriched television media. The practice allowed the formalization of some good practices for the design of additional content to the elderly user.

4 Observational Case Study

Aiming to investigate which format of additional content the elderly viewer feels more satisfied during an interaction with digital TV and to evaluate solutions that

make this interaction more flexible to meet the needs of different types of viewers, an observational case study was conducted with a group of elderly viewers. The main objective was to analyze the interaction of the elderly viewer with additional multimedia content offered to television programs and to identify which additional media format are more appropriate for this audience. The researchers started from the hypothesis that the additional content in formats of audio or animation, instead of text, make more sense and are best enjoyed by the elderly public and by the public with low education.

The case study was carried out in a center of social assistance (CRAS) in Brazil in a city with 200.000 inhabitants. The various actions undertaken by that center attract many people from various age groups. In particular, the center serves a group of elderly people aged between 60 and 85 years old, who play games and attend to physical and cultural activities. The goal of the partnership established between the service center and university is, in addition to promote the activities mentioned, to disseminate and promote the use of ICTs. In this sense, the case studies conducted with this group provided them access to devices such as smartphones and last generation television sets, including those with touch screen. The partnership was also a way to promote digital inclusion for these viewers.

The environment set up for this observation was in the CRAS, exploring an area that the target user attends. The environment was prepared similarly to a living room with TV, remote control and seats. Elderly citizens, from the group mentioned above, were invited to participate in an activity that consisted of watching a video enriched with additional content and use the remote to interact with such content. Twelve of them accepted to participate.

The elderly were separated into peers according to their age, education, psychomotor skills and experience with the remote. The intention was to group people with similar experience and profile and these data were collected from a profile questionnaire applied in the first meeting with this group. The additional content added to the video was in format of text, audio and animation. To evaluate the interaction of the elderly viewer with additional multimedia content in television programs, viewers were observed during watching sessions and interviewed in follow-up meetings for clarification of issues relevant to the research.

The genre of the media used was also chosen from data collected with the profile questionnaire. The data suggested that the target audience prefers to watch programs like soap opera and newscast. From this observation, a media displayed on a newscast was chosen to be enriched with additional content and to be used in the study. The media had about seven minutes and was enriched using Nested Context Language - NCL, a declarative language for authoring hypermedia. At each two minutes an additional content was inserted. These contents were designed to add useful and enlightening information for the audience and were inserted in intervals where there was not character's speech, in order to avoid damage to the main program. The first additional content was offered in audio format, the second in text format (see Figure 1a) and the third as an animation. The use of different formats allowed evaluating which of them is most appropriate for the elderly audience interaction with the additional multimedia content and which of them causes more satisfaction in this viewer.

During the ten seconds before the interval when additional content was inserted, an icon indicating the presence of this content ahead was displayed in the upper right hand corner of the TV screen (see Figure 1b). To trigger the additional content the elderly should press a proper button in the remote control during these ten seconds. The screen in which the media was displayed had 21.5 inches and the interactivity icon was a static interface element that filled about 2% of the screen.

Besides the profile questionnaire, were elaborated for this study a form to support the participants' observation and a Term of Agreement. The observation form was developed to analyze the interaction of the elderly viewer with additional content during the session, using the remote control. From this form it was possible to observe and report on whether the viewers observed the presence of the interactivity icon, if they used the remote control to interact and which content has been triggered by the remote control, if the viewers noted the presence of additional content after triggering it and, moreover, if the viewer interacted with the peer during the session. The Term of Agreement clarified to participants the objectives of the research, the risks and benefits and their freedom to quit their participation at any time. The activity was conducted following standards set by the Ethics in Research.

The second meeting with the elderly was divided into four parts. At the first moment the peers (previously grouped) were invited to watch the enriched video. At the beginning of the session, viewers were informed of how the session would be conducted and the importance of their participation. In the second part, an explanation about the elements involved in the session (mainstream media, interactivity icon, additional content and remote control) and the procedures to be followed to carry out the interaction with additional content through the remote control were explained. In the third part, each peer watched and interacted with the enriched video. The additional content was synchronized to the mainstream and there was not possibility to pause the video to pay attention to the additional content.

There were 2 remotes, one for each viewer. When the interactivity icon was shown, the viewer who wanted to see the added content could trigger it by pushing the green button on the remote see (Figure 1c). The viewers could interact with each other during the session.

Fig. 1. a) Additional content in text format. b) Interactivity icon. b) Viewers using remote control to interact.

In the fourth part, after the end of the media presentation, each viewer participated in an interview in which they were asked about their experience interacting with the additional content, which content attracted more attention and which content they considered more interesting and suitable to them.

5 Results

Among the six peers of elderly viewers who volunteered to participate in the case study there was 1 man and 11 women. Data about this audience collected from profile questionnaire show that most elderly live with some family member, but they have a tendency to watch TV alone. The data also show that 10 out of 12 elderly viewers use the remote to watch TV and they confess that the not understanding of parts of the program may affect their satisfaction.

The first additional content was presented during the first 2 minutes of the video and was a media in audio format. The second, appearing after 4 minutes of video, was a textual media presented superimposed on the video scene in yellow fonts and in accordance to guidelines and standards for Digital TV [16]. The third additional content, an animation, was presented after 6 minutes.

The data collected from the observation form show that approximately 10 out of 12 elderly viewers evaluated recognized, at least once, the presence of the icon of interactivity on the TV screen as being the interface element responsible for triggering the additional content, but only 3 elderly noticed the additional content offered thereafter. Many of them believed that the additional content was part of the main content. To preserve the main content authoring issues and considering that the awareness of the additional content is desired, this result suggests the need to provide additional contents in a more highlighted format and in a way that draw more attention. The 3 viewers, who noticed the presence of the icon of interactivity, triggered the remote control to interact with the additional content.

The data also show that 8 of elderly viewers interacted with their partner during the video presentation talking about the transmitted content. These findings reinforce the fact that television has a characteristic inherent, the collective audience.

At the end of the session, during the discussion time, all viewers could see the excerpts of the video with interactivity icons included and their related additional content. Hence, they could visualize the proposed solution by the research that aims to provide better appreciation of TV programs.

After watching these additional contents, viewers were asked about which of the contents they most enjoyed and which of them were more useful as a complement of the main program. The answers show the tendency of elderly to additional content in audio format, confirming our initial hypothesis.

Seven viewers chose the audio content, 1 viewer opted for the text content, 1 chose the content in the animation form and 3 said that all the contents were "*cool*". Among the viewers who opted for content in the audio format, were 4 illiterate ones. Of these 4, 3 prefer the content in audio format, showing that the viewers with less education prefer to interact with content in audio format. One viewer said: "*... I can't read, I prefer to hear ...*".

The viewer with higher education preferred the textual content. She stated: "*…I guess that the audio proposal does not disturb the video and it is cool, but I do prefer the text...*". This declaration indicates the need to offer flexible design solutions.

The viewer with higher education preferred the textual content. Observe the following excerpt from his declaration: "*…I guess that the audio proposal does not*

disturb the video and it is cool, but I do prefer the text...". This declaration indicates the need to offer flexible design solutions.

The following declarations reinforce the presented data about the trend of this viewers group for the content in audio format: "*...It is difficult to read...*"," *...The sound is easier to understand ...*". For the content in animation format, one viewer reveals "*…it attracted attention because there is something moving on the screen ...*".

6 Lessons Learned and Good Practices

Contributions of this study include lessons learned about the interaction of elderly people with additional multimedia content. The elderly is a potential public to the digital television audience and their interaction requirements should receive special attention from interactive television designers.

From the questionnaire and observation results, it was possible to verify that a static image filling only 2% of the TV screen does not appear to be sufficient to attract the attention of this viewers' group and, therefore, alert them about the availability of an additional content.

Most of the time, they were so watchful to the main video that they did not notice the presence of an icon for triggering additional content or they were so dispersed that they did not pay attention to the program or to the presence of an additional content. This situation can also occur with others profiles of viewers. It is necessary to consider another strategy to attract the attention of audience. A possible solution could be the use of a device for interaction that warns the viewer previously, for example, vibrating or beeping [17].

It could also be noted during the discussions that the additional animated content attracted the attention of some elderly viewers because, besides the audio, the moving on the screen indicated them that something different of the content and the main video was happening. This fact reinforces the hypothesis that a static interface element does not attract enough attention to elderly viewers. It was also observed that the content in textual format makes more sense to viewers with certain degree of education, because the textual content requires more ability of reading and comprehension.

From the data collected and observations made during the case study, and considering the characteristics of the elderly viewer, as described in Carmichael [8] in guide for development of interactive TV services, it was possible to formalize some good practices for the design of additional content for the elderly: a) The interactivity icon should be attractive, lively and preferably larger than 2% of the display screen; b) The interactivity icon should remain on screen for a time greater than five seconds, so that the user may become aware of its presence and trigger the remote control; c) Give preference to additional content in audio format or animation/audio; d) If possible, offer different options. Users with higher levels of education may prefer to interact with text; e) The additional content should be highlighted or have some strong evidence that it is not part of the original content. This would be useful for the elderly viewer to identify the difference between them and; f) It is necessary use a resource to

alerting the elderly viewer, since there is a risk of them do not pay attention to the icon of interactivity and miss the opportunity. One way to obtain attention and warn of the presence of additional content is through the vibration or beeping of the device, as suggested by Rodrigues et al. [17].

7 Conclusion

The case study allowed to evaluate the interaction of the elderly viewer with additional content, and also to observe which media formats are most appropriate for this group of viewers. By carrying out this study with diverse elderly viewers, different ages and education levels, the initial hypothesis that participants in older age and lower education level prefer the additional content in audio format was confirmed.

The proposal to draw attention to the additional content with vibration or beeping may be achieved by applications designed for mobile devices that already have such features.

Moreover, to make the additional content a feasible solution, new content professionals could be part of the production chain, complementing the content produced by TV broadcasters. Service providers could make content available to viewers via the Internet, for example.

This research also allowed that the elderly, who participate in the study, had the first contact and interaction with technologies such as Digital TV, promoting thus an initial step for the digital inclusion of those viewers.

Further studies may consider situations such as collective sessions, using different genres and media sizes, as well as different profiles of viewers should be conducted in order to identify other specificities and support the interaction design with additional content. Finally, an architecture that allows meeting the viewer's specificities, including the elderly public, will be proposed.

Acknowledgements. We thank CAPES for financial support, CNPq, FAPESP, FINEP and RNP that, through various projects of innovation and technological development, provided the physical resources for this research. Thanks also to the physical educator João Carlos, the Sao Carlos City Hall and the coordination of CRAS.

References

1. United Nations, http://www.un.org/esa/population/
2. Garcia, H.D.: The Third Age and the Internet: A Question for the New Millennium. Master's thesis submitted to Faculty of Sciences, Paulista State University, Marília – SP, Brazil (2002) (in Portuguese)
3. Reis, D.S.M.: Specifics of Communication for the Elderly. Journal of the Faculty of Communication FAAP 12, 65–71 (2004)

4. Colin, A.D., David, A.S., Wesley, K.T., Dilip, V.J.: Age, Affective Experience, and Television Use. American Journal of Preventive Medicine 39(2), 173–178 (2010), doi:10.1016/j.amepre.2010.03.020
5. Melo, M.A., Baranauskas, C.C.M.: Design for Inclusion: Challenges and Proposals. In: Proceedings of VII Brazilian Symposium on Human factors in Computing Systems, IHC 2006, Natal - RN, Brazil, pp. 11–20 (2006), doi:10.1145/1298023.1298026 (in Portuguese)
6. Fouts, G.T.: Television Use by the Elderly. Canadian Psychology 30(3), 568–577 (1989), doi:10.1037/h0079824
7. Davis, R.H., Edwards, A.E., Bartel, D.J., Martin, D.: Assessing Television Viewing Behavior of Older Adults. Journal of Broadcasting 20(1), 69–76 (1976), doi:10.1080/08838157609386376
8. Carmichael, A.: Style Guide for the Design of Interactive Television Services for Elderly Viewers. Independent Television Commission (ITC), Kings Worthy Court, Winchester, England (1999)
9. Brazilian Association of Technical Standards: NBR 15606-2: 2007, Digital Terrestrial Television - Data Coding and Transmission Specification for Digital Broadcasting, Part 1: Coding of data, corrected version - 07.04.2008, Brazil (2008)
10. Morris, S., Chaigneau, A.S.: Interactive TV Standards: A Guide to MHP, OCAP, and Java TV. Focal Press, Burlington (2005)
11. Rodrigues, K.R.H., Melo, E.L., Nakagawa, P.I., Teixeira, C.A.C.: Interaction with Additional Content to Support the Understanding of Television Programs. In: Proceedings of the IX Brazilian Symposium on Human Factors in Computing Systems, IHC 2010, Belo Horizonte -MG, Brazil, vol. 1, pp. 91–100 (2010) (in Portuguese)
12. Bulterman, D.C.A., Cesar, P., Jansen, A.J.: An Architecture for Viewer-side Enrichment of TV Content. In: Proceedings of the 14th Annual ACM International Conference on Multimedia, Santa Barbara, CA, USA, pp. 651–654. ACM (2006), doi:10.1145/1180639.1180775
13. Gradvohl, A.L.S., Iano, Y.: Matching Interactive TV and Hypervideo. IEEE Latin America Transactions 5(8), 579–584 (2007), doi:10.1109/T-LA.2007.4445709 (in Portuguese)
14. Teixeira, C.A.C., Linington, P.: Exploiting Interactive Video and Animation in Distributed Environment for the Design of Hypermedia and Graphical User Interfaces. In: Proceedings of VI Brazilian Symposium on Computer Graphics and Image Processing, SIBIGRAPI 1993, Recife - PE, Brazil, pp. 213–219 (1993)
15. DTV4All. Digital Television for All, http://www.psp-dtv4all.org
16. Instituto Nacional de Tecnologías de la Comunicación.: Digital Terrestrial Television (DTT) Accessibility Recommendations Guide, `http://www.inteco.es/file/snb-6ZR2I2DaXKiMJlkT_g`
17. Rodrigues, K.R.H., Pereira, S.S., Quinelato, L.G.G., Melo, E.L., Neris, V.P.A., Teixeira, C.A.C.: Interaction with Additional Content Using Multiple Devices to Support the Appreciation of Television Programs. In: Proceedings of Brazilian Symposium on Multimedia and the Web, WebMedia 2011, Porto Alegre – SC, Brazil (2011) (in Portuguese)

Exploring Children's Attitudes towards Static and Moving Humanoid Robots

Fang-Wu Tung[1] and Tsen-Yao Chang[2]

[1] Department of Industrial and Commercial Design, National Taiwan University of Science and Technology, 43,Sec. 4, Keelung Rd., Taipei, 106, Taiwan
[2] Department of Creative Design, National Yunlin University of Science and Technology, Yunlin County, 64002, Taiwan
fwtung@mail.ntust.edu.tw, changty8908@gmail.com

Abstract. This study investigates the uncanny valley for robots designed specifically for child users, and examines their attitudes toward humanoid robots with different anthropomorphic appearances and behaviors. An uncanny valley was identified in this study, indicating that the children were less attracted to images they considered highly human-like, although they were distinguishable from humans. The results of this study support Mori's uncanny valley hypothesis regarding children's perceptions of static robots. A significant finding is that moving robots can moderate an uncanny valley plot, which contradicts Mori's uncanny valley theory that the movement of a robot amplifies a generated emotional response. The moving robots exhibited various behaviors, such as facial expressions, speech, gazing, and gestures, which can generally enhance children's perceptions of robots. These behaviors were derived from human-human interactions, and can be considered social cues. The results of this study show that social cues can be applied to child-robot interactions. Children perceive robots are more socially and physically attractive when they exhibit sufficient social cues. Specifically, the display of social cues by robots that are less anthropomorphic can significantly enhance children's social perceptions of them. This has crucial implications for the behavior a child anticipates from a machine-like robot compared to a human-like robot. According to the study results, robots designed for children do not require excessively human-like designs. Middle- to low-level anthropomorphic designs combined with appropriate social cues can enhance children preferences and acceptance of robots. This also enables businesses to develop educational, care, or entertainment robots for children at a reasonable cost.

Keywords: humanoid robot, child-robot interaction, behavior, social cue.

1 Introduction

The paradigm of robotics is moving from a specific industrial technology to the consumer, home, and service markets. Service robots can be applied to people's living and working spaces as assistants, companions, or perform other social roles in the future [1]. Because robots are designed and built for purposes such as service, education, therapy, or entertainment, human-robot interactions become increasingly

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 237–245, 2013.

socially situated and multi-faceted. Social and emotional levels of interaction influence a person's acceptance of a robot's role [2,3]. Furthermore, considering a robot's functionality, previous studies have examined the social aspects of human-robot interaction to evaluate how people respond to the appearance and behavior of robots in various contexts [4,5,6]. To understand people's views toward robot companions in households, an investigation by Dautenhahn et al. [7] showed that the majority of people were in favor of a robot companion and saw the potential role of a robot as being an assistant, machine, or servant. People desired human-like communication with potential robot companions. Attributing human-like characteristics to robots with a human appearance is a design consideration because this could facilitate human-robot interaction [8]. Robots designed with anthropomorphic characteristics are the embodiment of a human-computer interface, and forms the basis for potential social relationships. This argument is supported by research by Goetz et al. [9] which indicates that people anticipate human-like robots would be best suited for interactive tasks, whereas mechanical-looking robots would be best suited for routine jobs. Scopelliti et al. [10] focused on people's attitudes toward domestic robots across three generations. Their findings showed that younger people scored higher on positive feelings (e.g., amusing, dynamic, pleasant, or relaxing) toward a domestic robot compared to adults and older adults people. Younger people who were born in the digital era did not express any anxiety toward the idea of a domestic robot. They reported a preference for robots with human-like attributes so they can interact with it in leisure situations, rather than perceiving it as a useful device. Considering children as potential users, the use of human-like features such as behavior and appearance can engage children to interact with robots. The large number of current humanoid robotics projects exemplifies this tendency.

As discussed, various humanoid robots are designed specifically for children, for educational, entertainment, and therapeutic purposes. Attributing human-like characteristics to a robot can facilitate children's understandings of its functionalities, and establishes meaningful human-robot interactions. Such notions lead to an assumption that increasing the realism of a robot has practical benefits. Although many researchers have pursued a highly human-like form of social robots, a potential danger is that highly human-like robots might fall into Mori's [11] uncanny valley. Uncanny valley theory hypothesizes a positive correlation between the human-like appearance and motion of robots with people's positive emotional reactions towards them. However, as likeness increases, there is a breaking point beyond which familiarity drops and robots become "eerie". When the emotional reaction is plotted against the robots' level of anthropomorphism, a negative valley becomes visible, and is commonly referred to as the uncanny valley. Furthermore, Mori argued that a robot's movement amplifies the emotional response in comparison to static robots.

The uncanny valley hypothesis has received empirical support from several studies. These studies have focused on examining Mori's hypothesis with adult participants, using static robot pictures as stimuli. However, whether the movement of a robot amplifies the emotional response compared to that of static robots has seldom been investigated. Robots are not developed to remain motionless, rather to be responsive to and interactive with people. Therefore, understanding people's perceptions of robots exhibiting human-like behavior provides insights that may be useful when designing a robot's appearance and behavior to facilitate better human-robot

interaction. Given that children are potential users of learning or entertainment robots, numerous humanoid robots have been created to serve as social companions or learning partners for children. This study investigates the uncanny valley for robots designed specifically for children, and examines their attitudes toward humanoid robots with various anthropomorphic appearances and behaviors. Obtaining insight about their perception toward robots that exhibit various degrees of anthropomorphic appearance and behavior could provide designers with valuable a reference

2 Method

2.1 Participants

For this study, we recruited a large sample group of 578 (N) children. We assigned 267 children to the static condition group, comprising 87 fourth graders (42 girls and 45 boys), 86 sixth graders (42 girls and 44 boys), and 94 eighth graders (45 girls and 49 boys). The remaining 311 children were assigned to the moving condition group, comprising 108 fourth graders (49 girls and 59 boys), 86 sixth graders (53 girls and 65 boys), and 94 eighth graders (43girls and 42 boys).

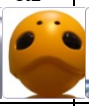
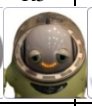
2.2 Experimental Stimuli

To conduct the investigation, we accessed numerous robot resources developed by companies, institutes, research labs, and artists, and acquired 54 robots ranging from "barely human" to "fully human". The 54 robot images obtained primarily from the Internet were edited by re-scaling and removing variables such as background color, marks, and other objects to ensure that the images were presented in a standardized format. Three professional designers subsequently examined the 54 images to remove redundant or inappropriate images regarding the anthropomorphism scale, and identified 34 images for inclusion in the study.

Twenty-nine children aged between 10 and 11 years were recruited to sort the 34 images by means of hierarchical clustering individual assessment. Upon commencement of a trial, the experimenters shuffled the cards in the set to ensure random presentation. To facilitate children's participation in this study, they were requested to sort the cards into the following three categories of realism: 1) low; 2) middle; and 3) high. Each subset was then divided into three groups based on identical criteria. Thus, each participant sorted the 34 robot images into nine groups ranging from low to high human likeness. Based on ranking data, we selected nine robots as experimental stimuli to show how children rate robots on an anthropomorphic scale, from "little resemblance to humans" to "highly human-like", (Table 1).

The robot pictures shown in Table 1 are the stimuli used in the static condition, whereas the robot videos are the stimuli employed for the moving condition. The length of the edited video clip is approximately 50 s, and the content fully demonstrated the behaviors of each robot, including locomotion, gesture, facial expression, and speech. Among the robots, R5 and R7 were capable of speech, whereas the remaining videos were accompanied by identical background music. Table 1 shows the behaviors of each stimulus presents in the video.

Table 1. The experimetal stimuli

Image / Behavior	R1	R2	R3	R4	R5	R6	R7	R8	R9
									
Speech					v		v		
Gesture	v	v	v			v	v	v	v
Facial expression					v		v		
Locomotion	v	v	v	v		v			v

2.3 Measurement Tools

The dependent variables for social attraction and physical attraction were modified from a version of McCroskey and McCain's [12] social and physical attraction scale, as well as from relevant studies that adopted the same scale to measure users' attitudes toward computers, robots, or media [13,14]. Social and physical attraction are two key dimensions of interpersonal attraction that have been found to facilitate interpersonal communication that leads to the formation of friendships. This study investigates whether various levels of anthropomorphic appearance and behavior influence children's social and physical attraction toward robots.

The social attraction scale comprises the following of five items: 1) I think this robot is friendly; 2) I like this robot; 3) I think this robot could be a friend of mine; 4) I would like to have a friendly chat with this robot; and 5) This robot would be pleasant to be with. The physical attraction scale comprises the following three items: 1) I think this robot is good looking; 2) I find this robot very physically attractive; and 3) I like the way this robot looks. We measured the two sets were measured using a set of paper-and-pencil questionnaires that employed a 7-point Likert scale ranging from 1 (very strongly disagree) to 7 (very strongly agree). The wording used in the questionnaires was designed after discussions with teachers and children to prevent any misunderstanding.

2.4 Procedure

Static condition group: the experiment was conducted using the discussed paper-and-pencil method. Each robot image was presented as a high-quality color printed image accompanying the aforementioned questionnaires. Participants were requested to evaluate the images of nine robots by completing questionnaires. The order of the stimuli was randomized for each child. The experiment was conducted at select schools. Participants completed the questionnaire survey in their classrooms or in a quiet place such as a school library.

Moving condition group: the experiment for the moving condition group was conducted using computers in a computer class room. The robot videos were played in random sequences. Each session contained 30 participants, and each participant was

provided earphones, and experimental personnel were present to assist the participants while watching the nine robot videos and completing in the questionnaire. The total duration of the experiment was approximately 40 min.

3 Results

We calculated the internal consistency (Cronbach's α) to assess the reliability of these scales. Cronbach's α results for the social and physical attraction items were all more than 0.7. According to Nunnally [15], Cronbach's α values of 0.7 are adequate for internal consistency and reliability. Therefore, the measures used in this study show adequate reliability.

3.1 Social Attraction

Table 2 shows the mean social attraction scores rated by participants in the static and moving condition groups. Robot R6 received the highest score for social attraction in the static condition group, whereas Robot R5 received the highest score for the moving condition group. The social attraction that children felt toward the stimuli is plotted in Fig. 1. Their attitude toward static robots supports Mori's prediction, and an uncanny valley emerged when children evaluated Robots R7 and R8. These robots were highly human-like in appearance, although they could be distinguished as non-human. However, this uncanny valley in the moving condition group was less apparent, which does not support the hypothesis that moving robots could amplify people's emotional responses compared to static robots. Generally, participants felt higher social attraction toward moving robots than those that were static.

Table 2. Mean Social attraction scores rated by Static and Moving condition groups

	R1	R2	R3	R4	R5	R6	R7	R8	R9
Static condition	3.31	3.30	3.40	4.30	4.49	5.57	3.76	2.62	4.45
Moving condition	4.89	4.72	4.74	4.81	5.71	5.26	4.85	4.12	5.03

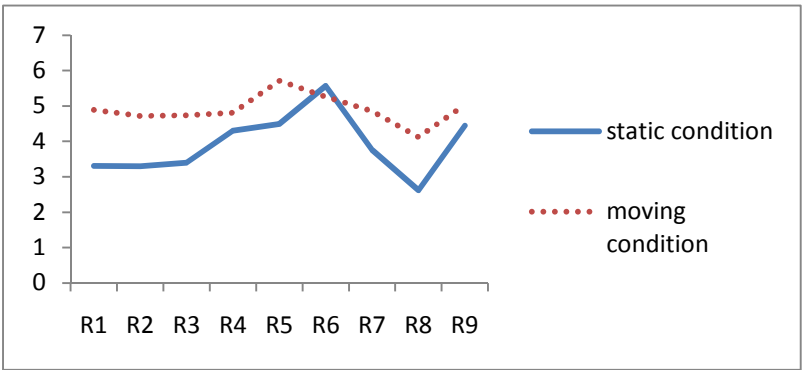


Fig. 1. Each robot's social attractiveness in the static and moving conditions

3.2 Physical Attraction

Table 3 shows the mean physical attraction scores rated by participants in the two conditions, and Fig. 2 shows the physical attraction children felt toward the static and moving robots. Similar findings were observed. Children reported that the moving robots were more attractive in appearance than those that were static. The curve is much flatter for the moving condition group than the static group.

Table 3. Mean Social attraction scores rated by Static and Moving condition groups

	R1	R2	R3	R4	R5	R6	R7	R8	R9
Static condition	3.03	2.85	2.96	3.75	3.76	5.09	3.07	2.37	4.4
Moving condition	4.52	4.39	4.28	4.49	5.01	4.91	4.04	4.02	4.9

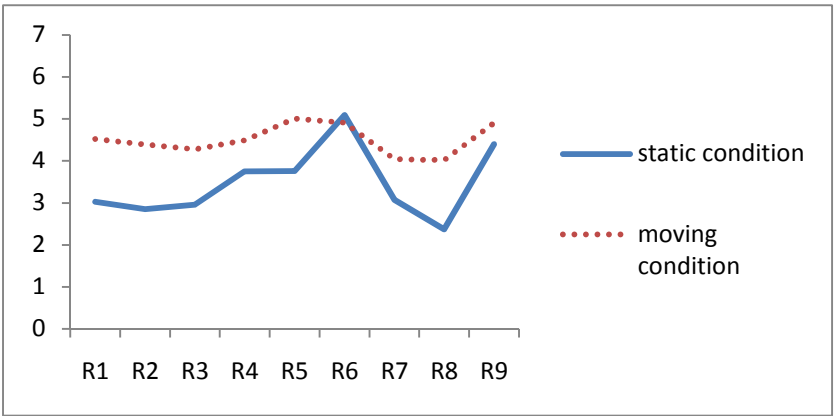


Fig. 2. Each robot's physical attractiveness in the static and moving conditions

4 Discussion

4.1 The Peak before the Uncanny Valley

The results from the static condition group show that the degrees of realism of robots have a significant influence on children's attitudes towards robots. Figures 1and 2 show a continuously positive change of realism versus children's evaluation of social and physical attraction until a point of realism beyond which children's evaluation decreases abruptly. As the appearance becomes less distinguishable from a human, children's responses become positive once again. Thus, an uncanny valley was observed in this study, indicating children were less attracted to images considered highly human-like, yet distinguishable from humans, which evoked a feeling of discomfort in the observer.

In the static condition group, Robot R6 received the highest evaluation from children for social and physical attractiveness. Robot R5 received the highest evaluation from the moving condition group. Robots R6 and R5 were at the peak before the

uncanny valley. This finding is identical to that obtained by Bartneck et. al [16], which shows that toy robots and humanoids preferred over humans.

An objective of this study is to identify the threshold of "humanness" that is appropriate for robots designed for children. The implication of the observation is that designers of robots intended for children might consider combining human and machine features to put effort toward the first peak of the plot, rather than attempting to perfectly replicate human-like appearance. The finding that children preferred Robots R5 and R6 among all stimuli supports Woods's [17] finding that children prefer robots with cartoon-like appearances. People, including children, are sensitive to the particular pattern of features that form a face. Using mere representations of cartoon-like faces can avoid the uncanny valley phenomenon and cover a large aesthetic range [18].

4.2 Movement Can Moderate the Uncanny Valley

A significant finding of this study is that moving robots can moderate the uncanny-valley plot. This contradicts Mori's uncanny valley theory which states that movement of a robot amplifies the emotional response more than a static robot. A comparison of children's attitudes toward static and moving robots shows a relatively flat curve for moving robots (Figures 1-2). Moving robots exhibiting behaviors such as facial expressions, speech, gaze, locomotion, and gestures can generally enhance children's perceptions of robots. The behaviors other than locomotion derived from human-human interaction can be considered social cues. Social cues (e.g., voice, presence of a face, and facial expressions) have been adopted into user interface design to enhance human-computer interaction. Studies related to computers as social actors (CASA) have empirically proven that people tend to treat a computer as a social entity when the computer adequately exhibits social cues to elicit social responses from people [19, 20]. CASA principles argue that computers which exhibit social cues can convey a sense of sociability and intimacy, thereby inducing social responses from people. This consequently improves user attitudes toward computers and fosters a more favorable relationship between users and computers. *Social Agency Theory* [21] also argues that a greater frequency of social cues in an interaction improves the quality of that interaction. The results of this study show that social agency theory can also be applied to child-robot interactions. Children perceive robots more socially and physically attractive when they exhibit more social cues.

Comparing the results of the static and moving condition groups, children gave higher evaluations to Robot R5 than Robot R6, which received the highest scores from the static condition group. In the videos provided in this study, Robot R5 could speak and show simple facial expressions with its eyelids and lips, whereas Robot R6 could only walk with whole-body motions. The fact that Robot R6 did not exhibit more social cues rather than locomotion could be the reason why it was the only robot to receive lower scores from the moving condition group than from the static condition group. According to Piaget [22], self-moving is a crucial criterion used by children to judge the "aliveness" of an object. Results for Robot R2 receive from both groups show that locomotion could enhance children's attitudes toward robots,

especially for those with low-degrees of human-like appearance. The more human-like robots appear, children might anticipate the robots to exhibit more human-like behaviors. In this sense, Robot R6, a humanoid robot, might be expected to have more social action capabilities rather than locomotion. This can also be observed in results for Robot R9, which has the appearance of an attractive young humanoid robot with influent whole-body motions. Robot R9 did not speak or show social cues such as facial expressions and gestures in the video provided. Thus, the moving Robot R9 did not significantly enhance children's social attraction toward it. This has significant implications for the accompanying behavior a person expects from a machine-like robot compared to a human-like robot. For instance, people might expect a human-like robot to have language capabilities, although they might not expect a machine-like robot to have any language communication abilities at all. Results from the moving condition group indicate that the human-like degree may not play an essential role in influencing children's emotional responses to robots. Appropriate action behaviors can enhance children's perception of humanoid robots, especially for robots with low- and mid-degree human-like appearances. This study recommends that designers consider applying social cues to robots to improve their social and physical attraction, in which child-friendly robot designs can be achieved at a reasonable cost.

5 Conclusion

In the innovation of robot service functions, developing educational or entertainment-oriented robots for children is a focal point for robot development. In this study, we investigated children's attitudes toward humanoid robots. Ultimately, by adopting this research as key reference, we hope to provide designers with design criteria for developing robots for children. According to the obtained results, robots designed for children do not require excessively human-like features. Mid- to low-level anthropomorphic designs combined with appropriate social cues can enhance children's preferences and acceptance of robots. The findings in this study also enable businesses to develop cost-effective educational, care, or entertainment robots for children.

Acknowledgments. This material is based upon work supported by the National Science Council of the Republic of China under grant NSC98-2221-E-239-012-MY2 and NSC100-2410-H-011-035-MY2.

References

1. Prats, M., Sanz, P.J., del Pobil, A.P., Martínez, E., Marín, R.: Towards multipurpose autonomous manipulation with the UJI service robot. Robotica 25, 245–256 (2007)
2. Bartneck, C., Forlizzi, J.: A design-centred framework for social human–robot interaction. In: Proceedings of the 13th IEEE International Workshop on Robot and Human Interactive Communication, pp. 591–594. IEEE Press, Kurashiki (2004)
3. Kiesler, S., Hinds, P.: Introduction to This Special Issue on Human-Robot Interaction. Human- Computer Interaction 19, 1–8 (2004)

4. Thrun, S., Beetz, M., Bennewitz, M., Burgard, W., Cremers, A.B., Dellaert, F.: Probablistic algorithms and the interactive museum tour-guide robot minerva. Journal of Robotics Research 19, 972–1000 (2000)
5. Falcone, E., Gockley, R., Porter, E., Nourbakhsh, I.: The personal rover project: the comprehensive design of a domestic personal robot. Robotics and Autonomous Systems 42, 245–258 (2003)
6. Kahn Jr., P.H., Friedman, B., Perez-Granados, D.R., Freier, N.G.: Robotic pets in the lives of preschool children. In: Poster presented at the 70th Anniversary of the Society for Research in Child Development-2003 Biennial Meeting, Tampa, FL, USA (2003)
7. Dutenhahn, K., Woods, S., Kaouri, C., Walters, M., Koay, K.L., Werry, I.: What is a robot companion–friend, assistant, or butler? In: IEEE/Robotics Society of Japan International Conference on Intelligent Robots and Systems (2005)
8. Friedman, B., Kahn Jr., P.H., Hagman, J.: Hardware companions? - What online AIBO discussion forums reveal about the human-robotic relationship. In: Cockton, G., Korhonen, P. (eds.) Proceedings of CHI 2003 Conference on Human Factors in Computing Systems, pp. 273–280. ACM Press, NY (2003)
9. Goetz, J., Kiesler, S., Powers, A.: Matching robot appearance and behavior to tasks to improve human-robot cooperation. In: Proceedings of the Twelfth IEEE International Workshop on Robot and Human Interactive Communication, Lisbon, Portugal (2003)
10. Scopelliti, M., Giuliani, M.V., D'Amico, A.M., Fornara, F.: If i had a robot at home Peoples'representation of domestic robots. In: Keates, S., Clarkson, J., Langdon, P., Robinson, P. (eds.) Designing a More Inclusive World, pp. 257–266. Springer, Cambridge (2004)
11. Mori, M.: The uncanny valley. Energy 7(4), 33–35 (1970)
12. McCroskey, J.C., McCain, T.A.: The measurement of interpersonal attraction. Speech Monographs 41, 261–266 (1974)
13. Moon, Y.: Similarity effects in human-computer interaction: Effects of user personality, computer personality, and user control on attraction and attributions of responsibility, Unpublished doctorial dissertation, Stanford University, Stanford (1996)
14. Lee, K.M., Peng, W., Yan, C., Jin, S.: Can Robots Manifest Personality? An Empirical Test of Personality Recognition, Social Responses, and Social Presence in Human-Robot Interaction. Journal of Communication 56, 754–772 (2006)
15. Nunnally, J.C.: Psychometric Theory. McGraw Hill, New York (1967)
16. Bartneck, C., Kanda, T., Ishiguro, H., Hagita, N.: Is the Uncanny Valley an Uncanny Cliff? In: Proceedings of the 16th IEEE International Symposium on Robot and Human Interactive Communication, pp. 368–373. IEEE Press, NJ (2007)
17. Woods, S.: Exploring the Design Space of Robots: Children's Perspectives. Interacting with Computers 18, 1390–1418 (2006)
18. Blow, M., Dautenhaun, K., Appleby, A., Nehaniv, C.L., Lee, D.: The art of designing robot faces: Dimensions for human-robot interaction. In: Proceedings of HRI 2006, pp. 331–332. ACM Press, New York (2006)
19. Nass, C., Steuer, J.: Anthropomorphism, Agency, and Ethopoeia: Computers as Social Actors. Human Communication Research 19(4), 504–527 (1993)
20. Reeves, B., Nass, C.: Media equation: how people treat computer, television, and new media like real people and places. Cambridge University Press, New York (1996)
21. Mayer, R.E., Sobko, K., Mautone, P.D.: Social cues in multimedia learning: role of speaker's voice. J. Educ. Psychol. 95, 419–425 (2003)
22. Piaget, J.: The child's conception of the world. Littlefield Adams, Savage (1929)

Part III

Interaction for Society and Community

Mobile Money Services in Uganda: Design Gaps and Recommendations

Rehema Baguma

Makerere University, School of Computing & Informatics Technology, Kampala, Uganda
rbaguma@cit.mak.ac.ug

Abstract. Mobile money is a great service for developing countries where the banking infrastructure is still severely limited and very few people can meet requirements of banking institutions. While these services are being greatly received due to the convenience that goes with electronic and mobile based services, the reception and uptake is still limited to literate populations who are the minority in these countries. This paper discusses the current design gaps of the mobile money service in Uganda based on a survey carried out in Kampala and Kayunga districts and provides recommendations on how these gaps can be addressed for the service to benefit more people especially the less literate poor that have no access or cannot afford conventional banking services.

1 Background

According to Duga and Getachew [2], the number of mobile phone subscribers is expected to hit 4.5 billion by 2012/2013 globally and this is attributed to emerging markets in Asia, Latin America and Africa. Africa alone has more than 300 million mobile phone subscribers according to the 2009 ITU report. In Uganda by June 2011, the mobile phone subscription had reached 14 million [12]. As mobile phones are multiplying in the developing world, new services are being innovated by Mobile Network Operators (MNOs) in addition to the traditional voice and SMS to remain competitive [9]. One of the prominent services is mobile money which loosely refers to money stored using the SIM (subscriber identity module) as an identifier as opposed to an account number in the conventional banking sense [11]. A notational equivalent in value is kept on the SIM within the mobile phone which is also used to transmit payment instructions. The corresponding cash value is physically held by the MNO, a bank or another third party depending on the business model [3]. MNOs and their agents provide an interface through cash-out (issuing cash on demand) or cash-in (convert cash into notational equivalent) functions providing convertibility between mobile money and cash (Morawczynski, 2009).

In Uganda, all the 5 major MNOs currently offer mobile money service namely: MTN Mobile Money from MTN, M-Sente from UTL, Airtel Money from Airtel, Warid Pesa from Warid and Orange money from Orange Telecom. According to bank of Uganda's supervision report for 2011, mobile money services registered much more significant growth compared to the previous year with the number of registered customers increasing from 1.7m in 2010 to 2.9m during 2011 [10]. MTN Mobile Money, the

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 249–258, 2013.

first of the five to launch and the biggest in the market, has reportedly registered more than 1,000,000 customers, setup over 1,500 agents/outlets across the country and transferred more than UGX 590 billion (US$ 245 million) since its launch in March 2009.

Like the case of mobile telephony versus fixed wire, Africa is expected to benefit more from mobile financial services because its financial services industry is not as developed as in Europe, Asia and North America [4]. However, while these services are being greatly received due to the convenience that goes with electronic and mobile based services, the reception and uptake is still limited to the literate urban populations who are the minority in these countries [11]. UNESCO [14] reported that one of the challenges of delivering mobile phone based services is that 41% of the population in developing countries is non-literate and even the literate among the poor are typically novice users of computer technologies. According to the 2011, Demographic and Health Survey (UDHS), the literacy rate of 15-24 year olds is 76.1% [13]. UDHS defines literacy as the 15-24 year olds who have attended secondary school or higher or who can read a whole sentence or part of a sentence. White [16] defines literacy for mobile phone users to mean the ability to understand the text that is displayed to them on the screen in order to be able to navigate the menu system, and then to understand the functionality that is subsequently presented to them. For technological literacy, White [16] defines it as the ability to understand or familiarity with common computer interaction paradigms such as the concept of menu systems, radio buttons or data entry. Chipchase [1] observed that non-literate populations avoid complex functions and primarily use phones for synchronous voice communication. UNESCO's Institute for Statistics defines the level of literacy as the percentage of people aged 15 and above who can with understanding, read and write a short, simple statement on their everyday life. Ndiwalana et al [11] noted that the bulk of mobile money transactions in 2010 in Uganda (including receiving) were happening in Kampala (the capital city of Uganda). In addition most respondents (72.6%) in Ndiwalana et al. study [11] reported having access to other financial services through a personal account in a formal financial institution and 72.7% reported being currently employed. This means that majority of the respondents in this study are not among the conventional mould of users expected to gain most advantage from using mobile money i.e. the unbanked rural population. White [16] noted that for businesses which are targeting poorer populations, ensuring that the products and services can be understood and used by non-literate people will increase their potential to succeed.

Therefore there is need for research on how mobile money service can be designed such that more low-literate and novice users of technology products who comprise the bulk of the unbanked in Uganda and other developing countries can take it up and use it. Medhi et al [8] revealed that non literate and novice users of technology face several barriers with existing text-based mobile interfaces such as difficulties understanding hierarchical structures, soft keys, scroll bars, non-numeric inputs, and specialized terminology. This paper discusses the current design gaps of the mobile money service in Uganda based on a survey carried out in Kampala and Kayunga districts and provides recommendations on how these gaps can be addressed for the service to benefit more people especially the less literate poor that have no access or cannot afford conventional banking services. The rest of the paper is organized as follows: methodology, design gaps for the mobile money service in Uganda, suggestions on how the design gaps can be addressed, and conclusion.

2 Methodology

To accomplish the objectives of this study, we conducted field interviews with current mobile money users and Agents (business people who buy franchises from MNOs to trade in the Mobile Money service) in peri-urban areas of Kampala city and rural locations in Kayunga district all in Uganda. All the participants owned or had access to a mobile phone. We interviewed 100 users (40 from Kayunga and 60 from Kampala), and 25 agents (10 from Kayunga and 15 from the suburbs of Kampala City namely; Najeera, Kiwatule, Natete and Kawempe. Kayunga District is a predominantly rural settlement in central Uganda. Interviewed users were met at mobile money shops and chosen randomly. The interview with users covered the languages spoken and written, functions on a mobile phone used, whether they find the mobile money interface easy to use, what is liked about mobile money, what they would change, value of mobile money to their day today activities, challenges faced using mobile money and possible solutions. The interview with mobile money Agents covered: the languages spoken and written, whether they find mobile money interface easy to operate, business viability of mobile money, what they like about mobile money, what they would change, desired versus available services, challenges faced and possible solutions. The obtained data was analysed using cross case/content analysis method to examine themes, similarities, and differences across the feedback obtained. The next section presents the findings from the study.

3 Findings from the Study

3.1 Social-Demographic Characteristics of the Participants

Of the 100 users interviewed, 60% were male and 40% female. 35 percent had a degree or diploma, 25% had completed Advanced Level (High School), 20% had completed Ordinary Level (Middle School), 15% had completed primary school (junior school) while 5% had no formal education. All the participants had ever used mobile money. They were residents of the area of study and 40% percent had personal bank accounts. Of the 25 agents, 80% were female and 20% male. The next section discusses the design gaps for the mobile money service in Uganda as per the feedback obtained from interviewing users and agents.

3.2 Design Gaps for the Mobile Money Service in Uganda

English Interface: The interfaces of all the mobile money services in Uganda is in English. This means the low literate masses cannot use it without intermediaries. Hinman and Matovu [4] noted that one of the main causes of people not engaging with mobile money services is due to illiteracy among other factors. From the interviews, users complained that customer care staff use English only and speak very fast. On the other hand, an Agent noted that most people from villages do not understand English hence they have to interpret for them but this keeps away some potential users who consider this an inconvenience or suspect that they can easily be cheated. Therefore the popular premise that mobile money through an increasingly large

mobile phone user base will provide a platform that could potentially be leveraged to service the financial needs of the poor is yet to be a reality. Additionally, even the literate but poor are not likely to use it or use it as much as they would due to association of English language with the wealthy and prestigious members of society. In a study conducted by Medhi *et al* [8], participants in the study on designing mobile interfaces for novice and low literacy users strongly and positively associated the English language with wealth and prestige. This was due to a combination of mindset inherited from colonial history as well as the modern-day fact of greater economic opportunities available to English speakers. The less educated that is those who cannot read and write in English at all or those with difficulty doing the same find it hard to learn and remember how to use mobile money. One user noted that "I stopped in primary five therefore I can only read and write in my local language-Luganda but mobile money is 'written' in English therefore I keep asking for help from Agents whenever I want to use the service". On the other hand, those who can read and write English without difficulty, reported that they found learning and remembering how to use mobile money services particularly the menu driven MTN mobile money very easy because of the prompts like one said "I follow prompts which is easy for anyone to follow if one can read English".

Limited Understanding of How Some Mobile Money Services Work: Services related to buying and selling goods or paying for utilities like water and electricity are still least used. One of the reasons for this is limited understanding of the m(e)-commerce concept as one respondent noted; *"I am still not sure that mobile money can reach when I use it to pay water bills otherwise it would be convenient"*.

Limited Coverage of the Payment Function: The payment function of mobile money service to-date covers a few aspects such as mobile phone credit, utility bills and school fees. However users are also using mobile money to settle informally other common payment obligations such as transport fares, professional fees, hotel/restaurant bills, rent, etc. However, the problem faced with this informal arrangement is that it depends on the willingness and flexibility of the recipient and in a majority of cases, they decline this form of payment due to various reasons such as having to bear the burden of converting mobile money into cash, limited understanding and trust of mobile money due to lack of sensitization, non-membership to mobile money services (transacting with a non-registered member attracts very high charges), failure to agree on transaction charges which the payee has to add on the bill, etc.

Desire for Mobile Money Service to Provide Services Provided by Banks: Some users view mobile money service as an alternative to conventional banks hence expect it to offer most of the services provided by banks. During the interviews, users requested MNOs to consider providing loans particularly airtime (mobile phone credit) and mobile money. They noted that this is particularly useful when one is in a remote place far from home. They reason that payment can be redeemed as soon as the user loads airtime/mobile money. At the moment, only Warid loans UGX 1000 airtime and none of the service providers loans mobile money. Viewing mobile money service as an alternative to banks is not surprising given that 100% of the respondents interviewed were all using sending, receiving and depositing money functions of mobile money services that are traditionally provided by banks.

Lack of Feedback on System Status during Interaction: For some services of mobile money such as paying payTV dues, users are not kept informed about what is going on during interaction. One user noted that: *"when you use a scratch card (an alternative option for paying TV dues), there is a person at the other end during the crediting process, so it is more reliable compared to mobile money where feedback is not guaranteed and in case of failure one is left in suspense concerning how far the transaction had gone"*.

Long Menu: Some of the users interviewed complained that the menu for mobile money services is very long hence difficult to remember how to use. MTN mobile money has 7 items at the main menu and between 1-4 options at sub menu level. Airtel Money also has 7 items at the main menu and between 1-4 options at sub menu level, M-sente has 8 items at the main menu and between 1-5 options at sub menu level while Warid Pesa is command driven. When asked if mobile money was easy to use, one interviewee responded*: "someone just took me through once but the menu is very long so it has taken me a lot of time to learn how to use it"*.

Confusing Functions: Users noted that there are some confusing functions such as redundant functions like non-mobile user function in MTN mobile money and the combined sending and receiving functions in Warid pesa. One commented that "the non-mobile user option under send money to in MTN mobile money confuses me for it serves no purpose". In Warid pesa, sending and depositing are not separated which confuses users. Users would like them to be separated like is the case with MTN mobile money.

Weak Error Prevention and Correction: The design of the current mobile money services has limited support for error prevention and correction in particular regarding sending money to the right recipients. According to the users and agents interviewed, MTN and Warid do not have adequate measures to protect against losses from errors or mistakes e.g. reversing a transaction or freezing an account when an error is made sending or depositing money on a phone. Many users interviewed reported having sent money to wrong recipients or received money from unknown people which could not be got back. Some of the responses given include:

> *"I have been receiving money from someone whose telephone number is almost similar to mine except for one digit and every time, I have to send it back because I am born again but I have heard of people who could not recover their huge sums of money because the people they sent it to did not cooperate"*.
> *"I Sent Money to a wrong Phone number and when I called the phone, the owner did not pick it"*
> *"I have never made a mistake. I am always careful when transacting but someone sent me UGX 900,000 last month which I had to go through the inconvenience of sending back"*.

One Agent called Musa in Kiwatule trading centre, in the process of sending UGX 250,000 had the receiver credited 3 times due to poor network connection. When this was reported to MTN's customer care, there response was that they could not do anything because the provision to block such numbers was removed since the process provides many levels of approval for the agent and customer to identify such anomalies

before completion of the process. In another incident at another service centre in Najeera run by Mariam Nansubuga, a customer ran away before paying a deposit of UGX 900,000 that had been processed. When MTN's customer care was contacted, they gave the same regret as in the first case. The other incident involved a fish monger called Mzee Zakayo who paid UGX 700,000 by mobile money to his supplier but sent it erroneously to a number that was switched off and the number was still off at the time of the interview. In desperation, the agents have resorted to informal negotiation with affected customers and other third parties to avoid mob justice and costly law suites. On the side of customers, they are losing confidence in the service particularly those that have been affected or have heard about the security incidents. Hinman and Matovu [4] noted that one of the reasons why some people are not yet engaging with mobile money services in Uganda is due to fear of losing money in the transaction among other factors.

Operation Tied to Unnecessary Constraints: For Airtel money, all transactions are limited to availability of airtime on one's phone yet charges do not come from airtime. Users suggest delinking transactions from availability of airtime on one's phone after all this is not where the charges are made.

System Interface Does Not Render Well on Computer: Agents of M-sente complained that accessing the system on a computer which some Agents prefer is a problem. That is it does not render seamlessly as it does on a phone. So the current interface of M-sente is device dependent which is a poor user interface design practice.

3.3 Other Gaps

Unstable Network. The network especially that of MTN is sometimes on and off like in the months of November 2011 to January 2012 and when it is off, business comes to a standstill hence loss of profits. This situation especially for MTN is attributed to a larger customer base compared to other MNOs. To make matters worse, there was no dedicated support line for mobile money Agents/customers for MTN, Airtel and Warid Pesa at the time of the interview, and the customer lines shared with voice services are always busy. Some of the responses from respondents include:

"When the network is off, we cannot work"
"Airtel's network is limited to towns- There is no network deep in the villages"
"You find people in villages wanting to use airtel money but network is a problem".
"UTL needs to improve network performance so that more people can use it.
"I am breaking even because I am dealing in other things i.e. it is not the only business I am involved in because availability of the network is limited to towns. Beyond, there is no network so people cannot use the service so we have very few customers"

Limited Sensitization/Education of the Public about the Service: Agents complained that service providers have not sensitized customers about the service which would create more demand hence more profitability.

Change of Terms without Consultation: Agents accused MNOs of regular change of terms of engagement without consulting them e.g. initially commissions on each transaction would be deposited on an Agent's account daily but at the time of the study, deposits were being made monthly which made it hard to track deposits. Agents and users would like to be more involved in the introduction of changes as key stakeholders.

3.4 Recommendations for Addressing the Design Gaps

Provide an Alternative Luganda Interface: MNOs need to consider providing a Luganda interface for mobile money in addition to the English interface as one of the strategies to increase usage among the less literate and poor sections of the population whose proficiency in English is limited. One Agent estimated that this could increase usage to more than 5%. Another Agent shared his opinion that a luganda version/option would make it easier for many to understand how the service works and how to use it. Some less educated users receive transactional messages much later after the transaction and have to spend time and money going back to the agent for interpretation. An option for the interface in a language they understand would save them money and time.

Consider Non Text Interfaces such as Spoken Input and Graphical Output: Medhi et al [8] in their study on designing mobile interfaces for novice and low literacy users established that text interfaces are unusable without literacy. They noted that non-literate subjects need non-text user interfaces such as spoken input and graphical output. Mobile money service providers should consider this option to cater for the user segment in the country that cannot read and write.

Visual Representation of Functions: Most people considered non-literate have some level of reading ability for example they can read numbers and can recognize at least a handful of symbols and words. But the design of most mobile services has not exploited this opportunity to make them more usable to them. To address this gap, White [16] recommends a visual icon based user interface that is in turn supported by voice prompts. The voice prompts help people navigate and learn the user interface as well as guide the users through the tasks supported by the service such as sending money. The prompts can also serve to provide confirmation of successful accomplishment of the given tasks.

Provide Mechanisms for Error Prevention and Correction at the Front End: According to Nielsen (20110), even better than good error messages is a careful design which prevents a problem from occurring in the first place. The mobile money design should incorporate effective error prevention and correction mechanisms such as options to suspend a transaction mid-way or immediately after (with in a given window period) in case of an error or fraud, blocking withdrawals for accounts pending security issues etc. Some MNOs already have some measures in place. For example, Airtel keeps message history and this function has made it more secure compared to others. In addition, during the sending process, the sender is given information about the name of the person he/she intends to send money to using registration information therefore in case of a mistake made one is able to know and correct it before proceeding. With M-sente during transacting, one is asked to enter a phone number twice and has to confirm the amount.

Encrypt Pin: Interfaces of all MNOs except Warid use encrypted pins. Warid should also implement use of encrypted pins during user login to prevent pin tapping from wrong elements.

Keep Users Aware of What Is Happening during Interaction: Nielsen (2012) in his Nielsen's 10 Usability Heuristics advises that a system should always keep users informed about what is going on. This can be achieved through appropriate feedback within reasonable time which Nielsen (2012) calls visibility of system status.

Use Fewer Menus and Dedicated Buttons: Dix et al (2009) noted that the short term memory capacity of a human being is limited and can only hold effectively 7± 2 chunks at a time. Hence the menu design of mobile money needs to keep the number of menu items at 7 maximum for a more usable interface. Jones et al. (2000) questioned suitability of menu-based navigation for novice users of computer technology which is the case with mobile money users and recommends designs with fewer menus and dedicated buttons for this target group. This view is supported by Lehrman [7]. Related to this, there is need to remove redundant functions and separate combined functions into distinct functions to make them easier to use.

Better Sensitization to the Public: MNOs should design and execute better sensitization campaigns for the public about the concept of mobile money to increase understanding of the service. This could partly be inbuilt within the service's interface. Hinman and Matovu [4] noted that the source of many people not engaging with mobile money is confusion with the mental model of the service. They observed that most promotional materials for mobile money services focus on awareness and benefits of the service. Hinman and Matovu [4] suggest that in order to fill the conceptual gap of how the service works, experience with creating and maintaining a formal account such as a bank account or account with a mobile service provider is required. The two noted that to the people in rural Uganda, the concept of money transfer is foreign but they do understand and have a wealth of experience with trading in particular exchanging assets that are equal in value e.g. exchanging a bull for a cow or a piece of land for a number of cows. Therefore it is such concepts that can be used in sensitizing the public rather than the abstract money transfer terminology. In Hinman and Matovu's [4] study, people who seemed to conceptually grasp how mobile money services work were the ones who related the service to buying and selling airtime and they noted that unlike transferring money, selling airtime mapped to their existing behavior of liquidating fixed assets when cash was needed or gifting fixed assets to people such as in the case of remittances.

Avoid Unnecessary Constraints in the Interaction Process: Airtel ties transactions to availability of airtime on one's phone yet this is not where the charges are made. This constraint and any others in this category that the study may not have found out should be removed to make the service easier to use for users.

Device Independence: Users of mobile money use a variety of devices ranging from high end to basic mobile phones as well as tablets, laptops and desktops. Therefore it is important to make the mobile money interface device independent such that it renders seamlessly on a variety of standard devices.

MNOs Should Share and Learn Each Other's Best Practices: instead of reinventing the wheel, MNOs should collaborate, learn from each other and put in practice each other's good practices. For example, some MNOs already have some effective error prevention and correction measures in place such as keeping message history and giving the sender information about the name of the person he/she intends to send money to using registration information such that in case of a mistake made one is able to know and correct it before proceeding by Airtel,. With M-sente, during transacting, one is asked to enter a phone number twice and has to confirm the amount. In addition, Agents have special dedicated numbers to call in case of a problem for quick help. These are measures MTN, Warid and Orange are lacking. Warid pesa and M-Sente could also learn and implement MTN's menu driven interface to replace their command line interface that users and agents find hard to use. In addition, Warid can learn from the rest the benefits of encrypting the pin during login and learn from them if necessary on how to implement it to provide more security to the users.

Other Design Considerations Suggested by Users and Agents:

- Current statements are too short e.g. MTN's mobile money statement covers the last 4 transactions only. Users recommended coverage of the last 10 transactions at minimum.
- Bonus: Users would like to be given a bonus for example when one sends/receives 500,000 Uganda shillings and above.
- Increase limit: Users would like an increment to the limit transferable in a day from 2,000,000 to 10,000,000 Uganda shillings.
- Provide for overdraft of a certain limit in times of emergencies: Users recommended an allowance for overdraft of at least Uganda shillings 50,000 to cater for emergencies
- Interoperability across networks: Users noted that the ability to transfer money between networks and banks would make the service more flexible to use for them. At the time of the study, only Warid had this service.
- Connection to Bank Accounts: Users would like to have their mobile money accounts connected to their bank accounts. This they said would help them link mobile money transactions with their bank transactions which would make overall management of their financial transactions easier.

4 Conclusion

Mobile money is a great service to developing countries where the banking infrastructure is still severely limited and very few people can meet requirements of banking institutions. The short time it has been in operation in Uganda has seen many achievements but also there are some challenges including design related that are threatening to stifle its usage among the people so much in need of it. This work has highlighted the design gaps and possible ways these can be addressed. We hope these insights will contribute to the growing body of knowledge around mobile money services and help to improve mobile financial services in Uganda as well as in other emerging markets.

5 Limitations

The work reported in this paper is based on factual findings from a group of respondents in the out skirts of Kampala city and Kayunga district. It is largely qualitative giving a description of the current design gaps of mobile money in the country according to the user experience of users and agents rather than a quantitative picture of users and Agent's opinions about the service. This approach was by design because the goal of the study did not require opinions but facts about the design gaps that could be verified independent of the participants.

References

1. Chipchase, J.: Understanding non-literacy as a barrier to mobile phone communication (2005), 789 hjyuv7v78hjyuv878yuhvjjjjjjj877vssxWebarekungarukamu juba (retrieved on May 16, 2012)
2. Duga, N., Getachew, H.: Mobile Services and ICT4D: To the Network Economy - Bridging the Digital Divide, Ethiopia's Case (UCC, 2010) (2012)
3. Esselaar, S., Stork, C., et al.: ICT Usage and its Impact on Profitability of SMEs in 13 African Countries. Info. Technologies & Inter. Development 4(1), 87–100 (2007)
4. Hinman, R., Matovu, J.: Opportunities and Challenges for Mobile-based Financial Services in Rural Uganda. In: CHI 2010, Atlanta, Georgia, USA, April 10-15 (2010)
5. Herzberg, A.: Payments and banking with mobile personal devices. Communications of the ACM 46(5), 53–58 (2003)
6. Hu, X., Li, W., et al.: Are Mobile Payment and Banking the Killer Apps for Mobile Commerce? In: 1st Hawaii Inter. Conf. on System Sciences. IEEE, Hawaii (2008)
7. Lehrman, S.: Dialing in. Mobile phones target the world's nonreading poor. Sci. Amer. 296(5), 30–31 (2007)
8. Medhi, et al.: Designing Mobile Interfaces for Novice and Low-Literacy Users. Designing mobile interfaces for novice and low-literacy users. ACM Trans. Comput.-Hum. Interact. 18(1), Article 2 (April 2011)
9. Mendes, S., Alampay, E., et al.: The innovative use of mobile applications in the Philippines-lessons for Africa. Sida Publications, Sida (2007)
10. Mugenyi, A.: Wanted: A Mobile Money Regulator. The CEO Magazine (June 2012)
11. Ndiwalana, A., Morawczynsk, O., Popov, O.: Mobile Money Use in Uganda: A Preliminary Study. In: Mobile for Development (M4D) Conference 2009 (2010)
12. Tentena, Mobile phone subscribers hit 14 million in Uganda. Business week June 20 , 2012 (2011), http://www.busiweek.com/index.php?option=com_content&view=article&id=1207:mobile-phone-subscribers-hit-14-million-in-uganda&catid=104:uganda& (retrieved on September 10)
13. Uganda Demographic Health Survey (UDHS) (2011), http://www.ubos.org/onlinefiles/uploads/ubos/UDHS/UDHS2011.pdf (retrived on June 20, 2012)
14. UNESCO, UNESCO Institute of Statistics Data Centre, regional literacy rates (2007), http://stats.uis.unesco.org/unesco/TableViewer/tableView.aspx?ReportId=y201 (retrieved on June 12, 2012)
15. UNICEF, Info by Country Statistics (2012), http://www.unicef.org/infobycountry/uganda_statistics.html (retrieved on June 16, 2012)
16. White, G.: Designing Mobile Services for non-literate communities. In: Proceedings of 2nd International Conference on M4D Mobile Communication Technology for Development (M4D) 2010, Kampala, Uganda (2010)

A Static and Dynamic Recommendations System for Best Practice Networks

Pierfrancesco Bellini, Ivan Bruno, Paolo Nesi, and Michela Paolucci

DISIT-DSI, Distributed Systems and Internet Technology Lab,
Dipartimento di Ingegneria dell'Informazione, DINFO,
Università degli Studi di Firenze, Firenze, Italy
{pbellini,ivanb,paolo.nesi,paolucci}@dsi.unifi.it

Abstract. Semantics computing technologies may be used to provide recommendations and stimulate user engagement in many kinds of services, such as social media, match making, best practice networks, technology transfer, etc. The recommendation metrics used take into account both static information and dynamical behaviors of users on a Social Network Platform. The recommendations provided include those realized taking into account also strategic and random users. The set of recommendations have been assessed with respect to the user's acceptance, which allowed to validate the solution and to tune the parameters. The experience performed in creating and validating recommendation systems adopted for ECLAP and APRETOscana best practice networks is described and results obtained are reported. The identified model has significantly increased the acceptance rate for the recommendation on ECLAP.

Keywords: best practice network, semantic computing, recommendations, social media, grid computing, validation model.

1 Introduction

Semantics computing technologies may be used to stimulate user engagement in many kinds of services, such as social media, match making, best practice networks, technology transfer, etc. The semantic computing is typically confined on the server side to provide recommendations. Despite to the massive success of social media, most solutions have limited semantic computing capabilities and provide simple recommendations about possible friends and on marginally similar content items. Among the possible combinations of suggestions related to users, content, ads, and groups only some of them are viable [1]. Recommendations should be computed on the basis of relationships U→U, G→C, C→U, etc. where U means User, G: Group and C: Content/Item, thus C→U means proposing Content suggestions to Users. The earliest solutions for guessing users' intentions have been based on keyword-based queries (i.e., sponsored search, or paid listing), which places ads and/or recommendations in the search results; and content match, also called content-targeted advertising or contextual advertising, which places ads on the basis of the web page content and content similarity [2], [3]. Contextual recommendations are widespread and many systems can extract keywords

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 259–268, 2013.

from web pages to produce suggestions [4], sometimes using semantic approaches [5]. In order to predict which terms describing a product or service are more relevant, models based on clustering, collaborative filtering, logistic regression, etc., are used [6]. User's ranking and reputation are connected to recommendations and trust and are becoming essential elements of web-based collaborative systems [7]. Implicit trust networks have been employed to incorporate trust and reputation [8], obtaining trust relations from a record of results in previous recommendations, by semantic reasoning and inference mechanisms upon recorded data.

In this paper, we reported the experience performed in creating and validating recommendation systems for services including complex descriptors: (i) ECLAP http://www.eclap.eu a best practice network and service derived from research tools and solution for providing services towards the community of performing art institutions; ECLAP includes about 120000 contents, 1900 users and 35 groups; (ii) APREToscana http://www.apretoscana.org a best practice service for supporting industries and research institution to match demand and offer and accessing to European commission founding; about 1800 users, 15 groups. They need recommendation systems for C→U, U→U, G→U, C→C in order to facilitate contacts. Contacts are consolidated by establishing stable connections among colleagues. Initially, both the above solutions where set up with a recommendation system developed for a medical best practice network, namely, Mobile Medicine [1]. The results obtained were not satisfactory since the context of ECLAP and APREToscana were quite different and thus a study phase has been started to reshape a more focused and tuned recommendation system.

To this end, a new model to compute similarities and propose suggestions has been developed. The proposed model to present suggestions has been validated. The validation aimed to verify if the modality and the parameters used to propose the recommendations where acceptable for the users in the domain. This model considers both the user profile and the information extracted by analyzing the actions that users perform on contents in the recent time. The list of recommended friends/colleagues it is not comprised only by the users they can more probably accept but is realized taking into account also strategic and random users, basing on the serendipity philosophy.

2 Requirements Overview

In this section, the main requirements for the recommendation system that can be adopted in best practice networks are discussed. The requirements have taken into account the lesson learnt from the management of a number of thematically different best practice networks, such as: ECLAP on performing arts, APREToscana research and technology transfer, Mobile Medicine medicine and emergency situations, IUF of CSAVRI (http://iuf.csavri.org) e-learning and new companies, etc.

The following requirements are referred to user to user (U→U) recommendation systems that have to provide reasonable suggestions to users basing on:

- both static and dynamic aspects of user behaviour, and of the content descriptors. User profile static aspects may include: age, languages, sex, city, job, education, preferred content, joined groups, etc;

- both new and regular users, avoiding the problem of the Cold Start. The new users often risk to do not have any recommendation since their static data are frequently not compiled and dynamic data are not yet collected;
- the last performed dynamic actions of the users and progressively forgetting the older (less recent) activities;
- the experience of the users that may have similar interests, intentions and/or temporal evolution in the portal;
- stimulating the connection to users in successive activities on the portal by creating side effects in their home pages (e.g., content posted, new connections, new groups). This action allows increasing the mean number of connection per user, and helps peripheral users to get connected with those with greater centrality, higher number of connections, etc.
- the complementary suggestions that could be unexpected for the recipient, but that can be accepted if well motivated (e.g., he likes this music genre, he visited Paris recently, etc.). So that, the system could learn from the acceptance of those new connections about the user preferences, despite the lack of related content and of similar colleagues;
- progressively estimate of the recommendations and not constraining the system to perform the systematic recomputation of all of them at each change of the dynamic aspects of the user profile and behaviour;
- an identification and use of the minimum number of parameters as a compromise from computation and effectiveness. This means that: (i) the computational complexity (the costs) of recommender system may be strongly influenced by the number of parameters taken into account; SVD/PCA and other statistical techniques can be used to reduce them as much as possible; (ii) the acceptance rate of a recommendation system may strongly depend on the presentation of the recommendations to the users.

3 Recommendations Model

As previously mentioned, recommendations can be computed through several different techniques. In most cases, the elementary operation is the similarity distance among descriptors. The estimation of distances among elements can be computationally expensive in the presence of complex descriptors and/or millions of items, depending on the complexity and on the high number of the descriptors. In the following, the model adopted in ECLAP is presented.

The **user static profile** consists of the data that change slowly over time [1]. In generalist social networks, the static profile is usually not very detailed: users do not like fill form online at the time of registration and sometimes tend to provide false information. In small thematic networks, however, this kind of information is much more reliable. The static profile takes into account: general information (name, surname, gender, date of birth, personal description, place of origin, spoken and mother languages), contact information (email and instant messaging contacts), school and work (school level, name of school or university, type of employment, name of the place of

work) and interest (a list of categories of interest). Some data are coded by using specific standard like: ISO 3166-1 alpha 2 and ISO 3166-2 for place of origin, ISO 639-1 for spoken languages. Type of employment and categories of interest depend on the portal usage domain. Moreover, the static profile considers also subscribed groups, friends list and user's interested taxonomy topics.

Table 1. Example of dynamic profile taxonomy based

User last N Interactions on	Corresponding collection of multilingual Taxonomy descriptor documents
Promoted	Arti figurative Danza Balletto, Arti figurative Danza Balletto Russi, Performing Arts Russian Ballet, Lettres Danser Ballet Russe…
Downloaded	Performing Arts Modern Dance Performance Utopy, Un altre Recerca Gènere Tema, Performing Arts Modern Dance Performance Utopy, Andre Forskning Genre Om, Andere Forschung Genre Gegenstand…
Played	Video Musicale Rock, Video Musical Rock, Video Music Rock, Video Musique Rock…
Favorites	Drama Gènere, Coreografia Rendiment Contemporani Dansa Arts Escèniques període històric Arts Escèniques, Performing Arts Modern Dance Performance Utopy, Drama Genre….

The **dynamic profile** is established on the basis of the actions the users perform on the portal. In ECLAP, the dynamic profile considers four types of user interaction: content seen by the user (played), user's favorite content (favorites), promoted and downloaded content. The user profile is built by providing a hierarchical taxonomical classification for each of the these categories considering the N last recently used content for each of them. Table 1 shows an example of a user that in the last period has watched several videos of rock music and promoted Russian ballets.

3.1 Users Proximity Evaluation

On the basis of the above mentioned aspects, the calculation of the proximity (*prox*) between two users A and B is defined by a linear combination of the values of proximity calculated for the static ($prox_s$) and dynamic ($prox_d$) profile aspects and defined as follows:

$$prox(A,B) = prox_d(A,B) \times \gamma_d + prox_s(A,B) \times \gamma_s \quad (1)$$

Where: γ_s , γ_d respectively weight the relevance of the static and dynamic distance.

$prox_s$ is defined by the eq. (2) as a function of static similarity between users relative to: spoken languages (v_{lang}), locality (v_{loc}), interests (v_{int}), common friends (v_f), subscribed groups (v_g), age (v_{age}), interested taxonomy topics (v_{tax}) as proposed in [1].

$$prox_s(A,B) = F\left(v_{lang}(A,B), v_{loc}(A,B), v_{int}(A,B), v_f(A,B), v_g(A,B), v_{age}(A,B), v_{tax}(A,B)\right) \quad (2)$$

$prox_d$The dynamic proximity is calculated using the value of similarity (*score*) provided by the system of indexing in Lucene and given by the Lucene's Pratical Scoring Formula in the multilingual documents created from document as described in Table 1. The dynamic profile is based on the multilingual document (doc_i) which is indexed. In this way, the dynamic proximity distance is defined as:

$$prox_d(A,B) = \frac{score(doc_A, doc_B)}{\max\limits_{X \in queryResult} score(doc_A, doc_X)} \quad (3)$$

The score provided is normalized using the maximum score given to documents resulting from the query.

3.2 Other Recommendation Criteria

The static and dynamic aspects satisfy a part of the requirements but not all. They are unsuitable to provide recommendations: (i) in the case of cold start (new users), (ii) that may stimulate new and unconnected users to get in contact to strong reputation colleagues, (iii) to strong users about new users that may need help in entering into the community. To this end, additional recommendation types have been added:

- **Strategic recommendations** are those that recommend to users who have a few colleagues those with highest number of connections and vice versa solving point (ii) and (iii).
- **Random recommendations** consist in suggesting to users a random selection of other users (perhaps with completely different to their own interests), driven by curiosity to new content, can create contacts with new friends, expanding the list of his interests and thus changing his dynamic profile.

As described in the following, the early validation presented in this document aimed at assessing the acceptance level of the ECLAP users about the proposed recommendations: static, dynamic, strategic, and random. The questions was: provided that the recommendations are performed by presenting some rationales about the similarity, which of the above mentioned aspects and recommendations would get the highest relevance from the user point of view, thus stimulating them to get in connection.

4 Computational Architecture

The architecture of recommendation system (Figure 1) consists of: the ECLAP portal, the ECLAP Storage Area and ECLAP Back Office, implemented by using AXMEDIS AXCP tools [9]. The ECLAP portal is responsible for: (i) collecting user static and dynamic information and storing all data respectively in the Static and Dynamic Data repository; (ii) creating users connections and store them in the User Relationship repository; (iii) providing suggestions to users and (iv) providing a survey to get a feedback by users for assessing and tuning the system. Finally, the ECLAP back-office is responsible of

evaluation recommendations according to the above presented model and executes on a distributed system multiprocessor algorithms:

- ***Potential Friends*:** it calculates the static proximity between users, performing the computations related to the new users, and renovating the computations for the less recently updated. This approach keeps the estimation as much light is possible to update the static aspects that slowly change over time.
- ***S.L.I.M.*** (Suggest Lucene Index Manager): it deals with the dynamic data by building dynamic profiles and indexing them in the Lucene engine. Also this process is periodic and estimates the new version of the documents (Table 1) indexing them incrementally, thus updating only the dynamic profiles of the most active users and those that have significantly changed their dynamic descriptors.
- ***US.TER*** (User Suggester): it calculates the vector of U→U recommendations by using the above mentioned method: (i) eq. (2) as described in the following, by considering value produced by *Potential Friends* and SLIM, (ii) random, and (iii) strategic.

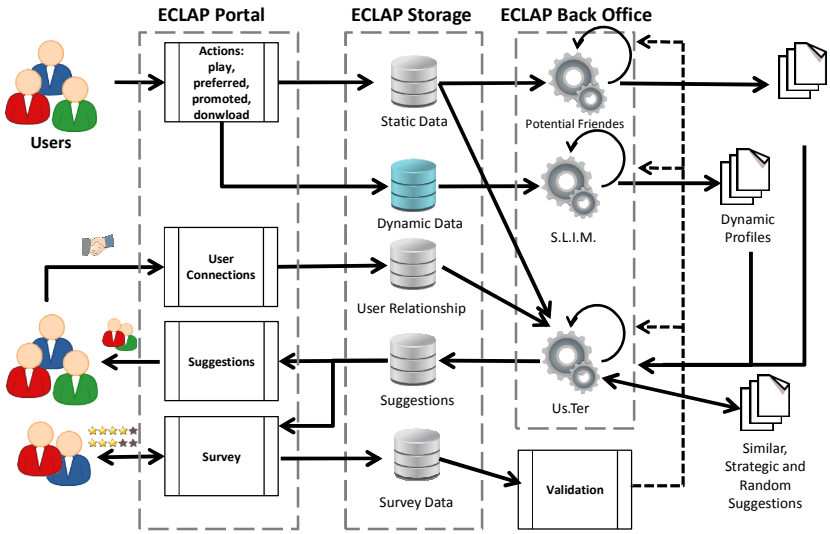


Fig. 1. Computational Architecture of the ECLAP Recommendation System

Typically, suggestions are provided with the 50% of kind (i), and 25% for kinds (ii) and (iii). The number of recommendations presented to the user have to be typically a small part of the whole recommendations computed, leaving at the users the possibility of taking more recommendations on demand.

The architecture is implemented as grid processes to take advantage from the distributed computing and to calculate/update progressively the needed data for generating recommendations. To have fresh suggestions, grid processes run as periodic processes according specific schedules defined by the administrator.

5 Validation and Results

Before moving to the real analysis of results achieved a system for validating the method. This allows tuning the system to match the user's preferences providing them better suggestions. A survey has been defined and posted on the portal. It proposes to each user the profile of 10 potential friends: each user is called to provide an answer to the question "*Are you interested in getting contact*?" by giving a vote (from 1 to 5) to indicate how much he is interested in the connection. To justify why the system asks to each user for (possible) potential friend, some motivations have been provided in the survey, as shown in Fig. 2. Such motivations have been built by considering:

a) ***similar users:*** similarities are shown according to the user profile as described;
b) ***strategic users:*** are motivated according to their activity or how much they are connected (or are not connected) to other users, etc.
c) ***random users:*** are randomly selected and motivated providing user profile details such as: the list of groups to which the users are registered, the profession, and the taxonomic classification of the last content viewed.

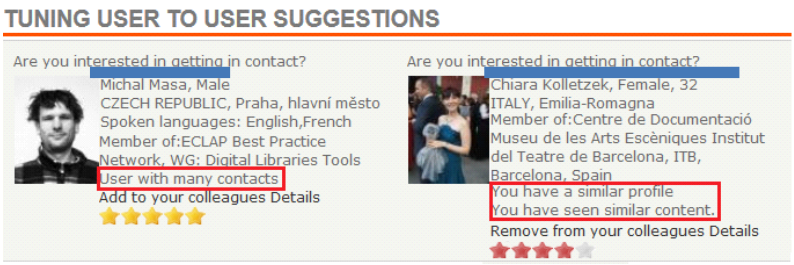


Fig. 2. Examples of the online survey (names have been obscured)

This should avoid that a user is faced with a recommendation without any details and thus considering it as not relevant. The range of the votes is 1-5. It is possible to consider the following three categories: *useful* ('Yes I want the user as my colleague!') if the vote is 4 or 5; *not useful* ('No I do not want the user as my colleague!') if the vote is 1 or 2; *not relevant* if the vote is 3 ('I do not know'). In order to estimate the model weights, a number of users have been involved into a learning phase in which we presented to them set of potential colleagues to be voted. Since the recommendations proposed were of two kinds, (a), (b) and (c) above, the validation votes have been taken into account in different manners, as reported in the following subsections.

5.1 Validation of Static, Dynamic Recommendations

This analysis aimed at estimating a linear model to the vote of users considering the values of metrics. More precisely, the model is defined by the following equation:

$$prox(A,B) = v_{lang}(A,B) \cdot \gamma_{lang} + v_{loc}(A,B) \cdot \gamma_{loc} + v_f(A,B) \cdot \gamma_f + v_g(A,B) \cdot \gamma_g + v_{age}(A,B) \cdot \gamma_{age} + v_{tax}(A,B) \cdot \gamma_{tax} + prox_d(A,B) \cdot \gamma_d \quad (4)$$

Where γ_i coefficient weights the relevance of the distinct proximity factors. The number of votes collected largely exceeded 10 times the number of γ_i weights of eq. (4). This allowed us to perform Multilinear Regression to estimate the weights as reported in Tables 2 and 3. Table 2 indicates that the estimated regression model has been confident: R_{square} represents the percentage variation in vote given by users explained by the model, and it is the square of R_m and it indicates the quality of the predictive model. F is the ratio of the variation of the votes that are explained and those that are not explained by the metrics: it is the variation of residuals. $F_{relevance}$ represents the probability that the vote can be random compared to the values of the metrics. In Table 3, γ_i weights, Stat-t and P-values of each metric are present. Stat-t is the ratio between the single metric coefficient inside the model and its standard deviation.

Table 2. Results on the data collected

Results (MR, first time)	
R_m	0.9624
R_{square}	0.9262
F	131.7795
$F_{relevance}$	2.3389E-33

Table 3. Multilinear Regression Coefficients analysis

Metrics	Coefficients (γ_i)	Stat-t	P-value
$prox_d$	-0.0047	-0.7312	0.4673
ν_{tax}	0.006	0.3906	0.6974
ν_{age}	0.0328	4.7163	1.3753E-05
ν_{lang}	0.032	10.2136	5.4563E-15
ν_{loc}	0.0408	6.0958	7.2905E-08
ν_{groups}	0.024	3.9658	0.0002

P-value indexes how the votes, compared to each metric, are relevant in the model: the higher is the Stat-t value and lower is the P-value. The analysis of the relevance of each metric used for the generation of suggestions (Table 3), reveals that the relevance of ν_{tax} and $prox_d$ are rather low. For this reason, a new Multilinear Regression has been realized and the results are reported in Table 4 and Table 5. The results highlighted that users selected their friends mainly on static aspects and less on the taxonomical modeling of the content. The data shows that the motivations that drive users to tighten social connections are: age, spoken language and location.

Table 4. Results on the data collected

Results (MR, second time)	
R_m	0.9620
R_{square}	0.9254
F	201.4967
$F_{relevance}$	1.6585E-35

Table 5. Multilinear Regression Coefficients analysis

Metrics	Coefficients (γ_i)	Stat-t	P-value
ν_{age}	0.0324	5.0306	4.1064E-06
ν_{lang}	0.0154	11.6651	1.3683E-17
ν_{loc}	0.0131	6.1930	4.4967E-08
ν_{groups}	0.0027	4.4234	3.7840E-05

5.2 Validation of Other Kinds of Recommendations

On the basis of the votes collected, the strategic suggestions have been considered very interesting and useful by users, that in the 73,81% of time admitted to have been

convinced of getting in connection with the proposed users. On the other hand, randomly provided users did not give a real stimulus to get connected. In fact, in this latter case, the percentage of votes in the three categories have been very similar: *useful* 32,58%; *not relevant* 37,08%; *not useful* 30,04%.

5.3 Impact of the Produced Model on Acceptance Rate of Recommendations

Once discovered these issues, the derived model has been adopted in ECLAP, substituting the previous model that was in place since 28 months. Therefore, from November 2012 up to now, the new model has been adopted, and use data have been collected, in order to assess the users' appreciation. From the data analysis, it can be noted that the increment of the average number of accepted recommendations is of the 42%: ECLAP registered an averaged increment of 42% of the accepted recommendations per week in the period Nov 2012 - Feb 2013. Moreover, by comparing the same periods: 'Nov 2011-Mar 2012' (old model) against 'Nov 2012-Mar 2013' (new model), an increment of connections of 241% has been registered.

6 Conclusions

In this paper a recommendation system integrated in a collaborative best practice portal has been described. The system provides to users a list of potential colleagues based on both static information and dynamical behaviors of users. These aspects have been used to define metrics that combined together allow to estimate a proximity assessment between two users. The model includes a set of weights to define the relevance of the different metrics. The acceptance rate of the recommendation is not only an aspect related to user proximity but also the manner by which the recommendations are proposed. In this work, we have validated the model to propose different kinds of recommendations: static/dynamic based, random and strategic. The assessment performed allowed us to tune the recommendation system increasing the previous solution. The analysis reveals that the system has been useful to increase the number of acceptance rate of the new connections and suggested the actions to be performed in order to improve the recommendation system efficiency.

Acknowledgement. The authors want to thank Dario Nesi, Lorenzo Calosci for their collaborations in the implementation of some of the tools and all the single people and partners involved in ECLAP, and the European Commission for funding in the Theme CIP-ICT-PSP.2009.2.2, Grant Agreement No. 250481.

References

1. Bellini, P., Bruno, I., Cenni, D., Fuzier, A., Nesi, P., Paolucci, M.: Mobile Medicine: Semantic Computing Management for Health Care Applications on Desktop and Mobile Devices. International Journal Multimedia Tools and Applications 58(1), 41–79 (2012)

2. Ciaramita, M., Murdock, V., Plachouras, V.: Semantic Associations for Contextual Advertising. Journal of Electronic Commerce Research 9(1) (2008)
3. Anagnostopoulos, A., Broder, A.Z., Gabrilovich, E., Josifovski, V., Riedel, L.: Just-in-Time Contextual Advertising. In: CIKM 2007, Lisboa, Portugal, November 6-8 (2007)
4. Zhang, Y., Surendran, A.C., Platt, J.C., Narasimhan, M.: Learning from Multitopic Web Documents for Contextual Advertisement. In: KDD 2008, Las Vegas, Nevada, USA, August 24-27 (2008)
5. Broder, A., Fontoura, M., Josifovski, V., Riedel, L.: A Semantic Approach to Contextual Advertising. In: Proc. of 30th Annual Int. ACM SIGIR Conf. on Research and Development in Information Retrieval, SIGIR 2007 (2007)
6. Bartz, K., Murthi, V., Sebastian, S.: Logistic Regression and Collaborative Filtering for Sponsored Search Term Recommendation. In: EC 2006, Ann Arbor, Michigan, June 11-15 (2006)
7. Anderson, E.F., McLoughlin, L., Liarokapis, F., Peters, C., Petridis, P., de Freitas, S.: Developing Serious Games for Cultural Heritage: A State-of-the-Art Review. Virtual Reality 14(4) (2010)
8. Martín-Vicente, M.I., Gil-Solla, A., Ramos-Cabrer, M., Blanco-Fernández, Y., López-Nores, M.: Semantic inference of user's reputation and expertise to improve collaborative recommendations. Journal of Expert Systems with Applications: An International Journal Archive 39(9), 8248–8258 (2012)
9. Bellini, P., Bruno, I., Cenni, D., Nesi, P.: Micro grids for scalable media computing and intelligence on distributed scenarios. IEEE Multimedia 19(2), 69–79 (2012)

Connecting Electric Vehicles and Green Energy

Peter Dempster

BMW Group Technology Office USA
Peter.dempster@bmw.de

Abstract. This paper discusses the interrelationship between the purchase of green energy (GE) and electric vehicles (EV) and the motivations for and values formed around the purchase of the combination of the two. The BMW Group completed a two-year EV and GE user study, a joint project with the PH&EV Research Center at UCDavis and multiple market research agencies. Through a focus group study of US East and West coast all-electric MINI E markets and a new car buyer online survey, the BMW Group assessed the value of offering GE with EVs. A follow-on user experience survey, assessed the success of two GE products in the US. 39% of MINI E and 29% of Active E respondents have purchased home solar. A vehicle design game revealed that adding GE options increased overall demand for EVs among conventional buyers by 23%, with technology interest, environment and cost savings reported as the motivation.

Keywords: Electric Vehicle, Green Energy, Solar, Environment.

1 Electric Vehicles Deliver a Positive User Experience

Vehicles with plug-in electric drive technologies deliver a positive user experience, according to published customer satisfaction surveys. More than 90% of BMW Active E, Nissan Leaf and GM Volt customers report being satisfied or very satisfied with their electric vehicle purchase. After two years with the car, 88% of respondents of the MINI E survey said that they would purchase an EV within the five years following the field trial. After one year with the BMW Active E, 91% of users responded "yes" to the question; "Do you plan on buying or leasing an electric vehicle in the next three years?" according to preliminary data from a recent survey.

EV customers report numerous advantages over vehicles with conventional drive trains. These include a quiet interior space, smooth and quick driving experience, compelling environmental benefits, the ability to refuel at home, increased energy independence and decreased fuel and operating costs. EV drivers form new values such as the intersection of clean and fun-to-drive, the ability to expand an electric driving territory and to master energy use, in and out of the car [1]. EV users discover that they can be independent from an energy system which is potentially unsustainable and are able use electricity from a portfolio of clean generation technologies, including GE, to fuel their car.

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 269–276, 2013.

2 BMW Group Electric Vehicle History

BMW Group has introduced 7 electric vehicle models since 1972, but only two EV field trials have been coupled with academic user research. The first, an all-electric MINI E field trial occurred from 2009 to 2011 with 450 vehicles in regular customer hands on U.S. East and West coast markets. The focus of the accompanying user study - which included multiple in-person and phone interviews, online surveys, focus groups and vehicle data logging - was driving and charging behavior, gauging the interest of the market in electric drive technologies and renewable energy, range preference and demands on the electricity grid.

BMW of North America is currently leasing 700 BMW Active E vehicles to customers in Boston, Hartford, New York City, New Jersey, Los Angeles, Sacramento, San Diego and San Francisco for a period of 24 months. The BMW Active E is unique compared to past BMW all-electric offerings, like the MINI E. First, the active E is equipped with a standard charging port (J1772), so customers can use tens-of-thousands of public charging stations which greatly enhances everyday usability. Second, Active E customers can purchase home solar through a preferred partnership with Real Goods Solar or a wind power renewable energy certificate package through Green Mountain Energy. Third, the Active E has an advanced telematics system, ConnectedDrive, and a smart phone application, MyBMWRemote, which allows the customer to connect to the car to see its status, like battery state-of-charge and remaining range, and set charging and preconditioning timers.

3 EV Charging and the Environment

The environmental benefits of driving an EV are linked to the timing, duration and location of charging. For example researchers at the Institute for Transportation Studies at UCDavis identified through an electricity supply model that emissions associated with generation vary greatly with hourly demand and power plant availability [2]. Green house gas (GHG) emissions associated with electricity generation also vary greatly across geographic regions. For example, the US South, Midwest, and Western Rocky states have more than 3 times greater EV GHG emission rates than the West Coast and Northeast [3]. An electric vehicle charged in California is likely to produce 60% less well-to-wheel GHG emissions than a conventional gasoline vehicle [4] because of the mix of electricity generation sources. In 45% of US regions, regardless of when it's charged, an EV produces less well-to-wheel GHG emissions than even the cleanest hybrid vehicles [5]. The BMW i-3 is expected to achieve a 50% reduction in well-to-wheel greenhouse gas emissions largely due to the integration of GE in both production and use phases of the vehicle.

When charged intelligently, electric vehicles can be grid assets, lowering electricity costs for all consumers and total emissions from the electricity sector. A fleet of plug-in vehicles can act as: 1. a virtual power plant, eliminating the need to turn on expensive natural gas power plants and reducing costly transmission and distribution system upgrades, which will lower the cost of electricity for all consumers, and 2. a large battery,

storing wind energy which may be produced in excess during the night when load is at its minimum or other renewable energy which is intermittent so that it can be used later, which will lower total grid emissions and enable the installation of more GE.

An ongoing BMW Group partnership with a major European energy service provider, Vattenfall, is demonstrating that intelligent charging of an EV can reduce costs and GHG emissions associated with EV charging. In a Berlin field trial, a cloud computing algorithm calculates and sends a charging profile to the MINI E based on a wind energy forecast, a customer mobility calendar, and battery state-of-charge and state-of-health information stored on BMW backend servers. Users are encouraged to plug in according to their mobility calendar through gamification of the smart phone application. Through intelligent charging of the MINI E, GHG emissions are reduced by 3%. Further economic and environmental benefits can be realized through a true vehicle-to-grid scenario; Excess wind electricity is purchased during off-peak hours and fed back into the grid during peak periods when electricity is often more expensive and polluting.

The BMW Technology Office USA, located in Mountain View CA, recently launched eMobility Lab, a new research and demonstration platform for the evolving ecosystem surrounding sustainable electric mobility. eMobility Lab is focused on five

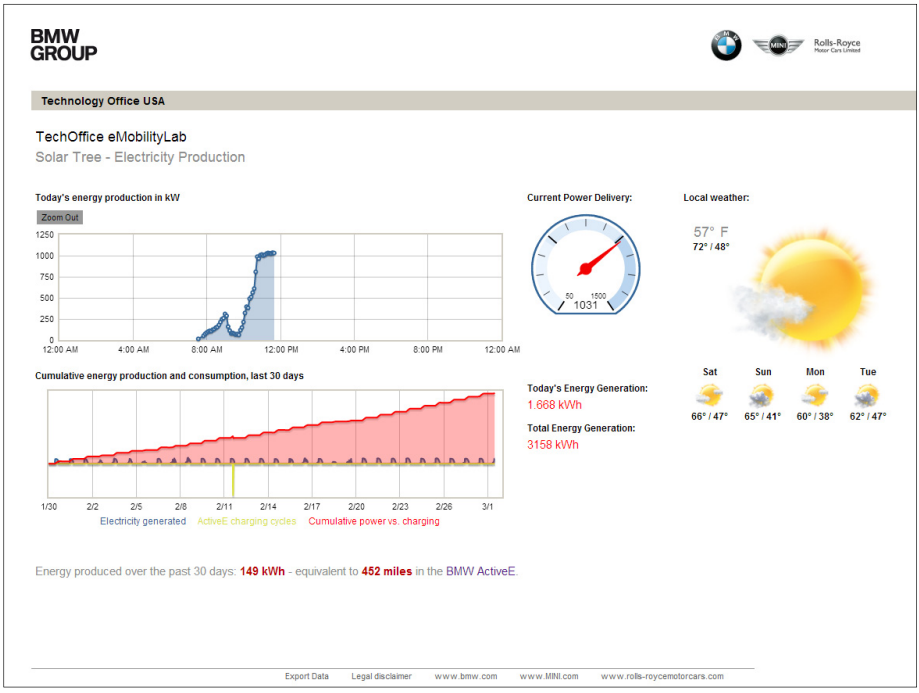


Fig. 1. BMW Group Technology Office USA solar and EV energy visualization tool

key research areas: 1. understanding the interaction between renewable sources of electricity and EV charging (Fig. 1), 2. communication pathways that allow the consumer and utility to visualize and interact with vehicle charging and home energy use, the "smart home", 3. second-life opportunities for EV batteries, 4. consumer studies and 5. electric vehicle grid services. BMW Technology Office USA is keeping an eye on emerging trends in technology innovation and learning from consumers about emerging markets through focus group and survey based studies.

4 Green Energy and EV Users

A survey of customers who participated in a two-year all electric MINI E US field trial revealed a strong user preference for charging EVs with solar (100%) and wind (98%) energy. The same survey of former US MINI E owners found that 37% owned solar by the end of the field trial. This result was re-confirmed by a recent survey of California EV owners (96% Nissan Leaf owners) which found that 39% have invested in home solar systems (n=1419), helping to "fuel" their vehicles with renewable solar energy [6]. Most recently, 29% of Active E respondents said that they have home solar according to preliminary results of a recent survey. An additional 25% responded that they intend to purchase solar within the next three years.

In order to gain a better understanding of the connections that EV drivers make with GE and the motivations for combining the two, BMW Group and the Plug-in Hybrid and Electric Vehicle Research Center at the Institute for Transportation Studies at UCDavis conducted eight focus groups in Beverly Hills, CA, Manhattan, NY, and Montvale, NJ in June of 2011. Participants were drawn from three sets of households: all-electric MINI E customers, people who bought GE through their local utility (ex. LADWP Green Power Program), and people who owned conventionally-fueled MINI vehicles and who had no obvious connection to GE or electric mobility.

By creating a comfortable environment in a focus group, the moderator enabled MINI E customers to tell a story about their common experience with owning, driving and charging an EV. The discussion began with MINI E users describing how they came to be an EV owner - how the MINI E made them feel at the start of the field trial along with initial like and dislikes - and moved to how they currently see the MINI E, electric mobility and fueling an EV. The moderator then asked if MINI E users knew about GE options available to them and if talking about GE enhanced their feelings about and experience with the MINI E. The group discussion was conducted four times in three locations so that the research team could identify themes and patterns. These themes were incorporated into the design of the survey component of the mixed-method GE and EV buyers study.

During the focus groups MINI E customer response to the combined offering of GE and EVs was truly mixed. Some participants had already made large investments in home solar photovoltaic (PV) systems or were participating in voluntary GE programs. Leasing the EV was an extension of a sustainable lifestyle which included GE.

I put SunPower panels on my barn so I have a 20-kilowatt solar system and the MINI E was kind of a segue to that.

Some customers for whom home solar was not accessible had thought about other innovative pathways for coupling solar or wind power and EV charging, like a power purchase agreement or intelligent charging.

I can buy a package of green power, renewable power, power generated from renewable sources that [makes me] feel good that a very healthy portion of my driving - if I didn't have solar - is covered by a program that's really easy for me to sign up on or with to drive basically in a non-emission-generating way.

A third group decided that because EVs were so clean, it was not necessary to charge them with GE. They believed that regardless of how it's produced, using coal or other sources, electricity must have a lower environmental impact than gasoline burned in an engine.

It would have to be a really compelling way in which you packaged whatever it is...I already feel great about driving the electric car. I don't need to feel great about paying Southern California Edison more money for the proper source, for me, just me.

The nationwide new car buyer online survey applied themes and language taken from the focus group sessions to assess consumer interest in PEVs and green electricity programs. Deployed in July of 2011, the survey contained an EV and GE design game. BMW Group and a market research agency recruited three U.S. samples to complete a web-based survey, including recent buyers of new conventional vehicles (CVB, n=1064), hybrid vehicles (HEVB, n=364) and plug-in electric vehicles, including the MINI E, Chevrolet Volt and Nissan Leaf (PEVB, n=74).

Respondents were asked to design a vehicle, choosing the make, model and engine type (conventional, hybrid, plug-in hybrid or electric). They then designed a home energy plan, choosing between no green program (or current "green" program if already enrolled), a monthly green program applied to at least 20 percent of home electricity use with a user defined electricity source, a 2-year green lease supporting wind or solar power through the voluntary purchase of renewable energy certificates (and requiring a two-year commitment) and purchasing and installing residential solar, financed as one monthly bill for 20 years (a power purchase agreement or PPA). The final section of the survey asked them to choose a vehicle and electricity plan side-by-side with an option to cover EV electricity use with GE. Respondents were asked to identify motivation for choices of vehicle, energy plan and combinations of the two.

Respondents that already have "green electricity" are more likely to design PEV in the game. Most hybrid and plug-in buyers design solar (18-35%), or already own solar (32-37%). Adding GE options increased overall demand for PEV designs among CVBs by 23%, HEVBs by 20%, and among PEVBs by 5% [7]. For PEVBs the most frequently designed combination (38%) is an EV with home solar; 86% combined an EV and GE and none designed a conventional vehicle and rejected GE (Fig. 2).

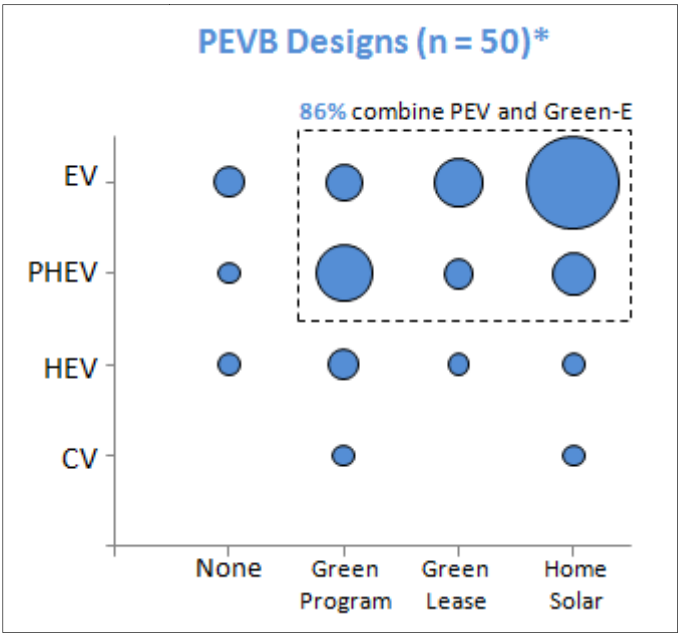


Fig. 2. EV buyers combine vehicle and energy plan (source: [7])

Motivations for combining GE and EVs vary within car buying groups. They include environment, total cost of ownership, energy independence, renewable energy support and control over electricity sources. Conventional and hybrid vehicle buyers are more likely to be motivated by cost savings. EV buyers are more strongly motivated by technology interest and as well as environmental benefits. Customers' reported motivations provide guidance on policy and marketing strategies to advance GE and EV demand.

5 Green E Product Offerings

In 2012 the BMW of North America offered Active E customers two GE options. One, a wind-energy renewable energy certificate (REC) package offered through Green Mountain Energy, provides a two-year green energy plan - approximately $48 - that corresponds to the amount of energy an Active E uses during the lease, 6400 kWh, with a mechanism for "topping up" at the end (Fig. 3). The second is a preferred partnership with Real Goods for a residential solar PV installation which is appropriately sized for home and EV. A portfolio approach to green energy allows BMW Group to meet all US EV customer needs.

The most recent results from the Active E user experience survey show an encouraging take-rate for BMW GE products offerings. The largest customer of the Green Mountain Energy REC package is DriveNow, a 70 EV car sharing program in

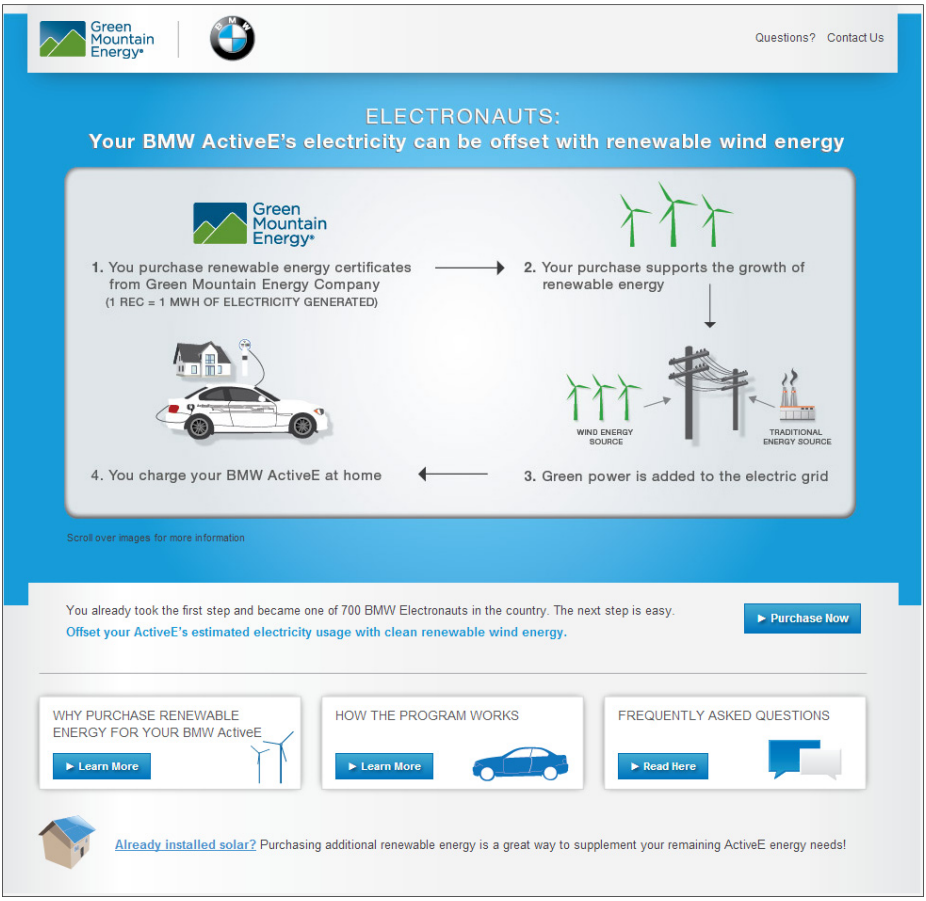


Fig. 3. Active E customers can purchase RECs through Green Mountain Energy

San Francisco. GE gives DriveNow a competitive advantage over other EV car sharing programs and is expected to attract environmentally minded consumers, which is why the GE program is front-and-center in marketing materials for the company. According to the preliminary results of the Active E user experience survey, 5% of respondents have purchased home solar through Real Goods and 3% of respondents have purchased wind energy through Green Mountain Energy.

6 Conclusions

In a multi-year project exploring the consumer demand for EVs and GE, BMW Group and numerous research partners completed a novel study, in three parts. The first, a literature review, aided the design of a unique multi-methodology study employing both focus group and survey research. The second, focus groups with MINI E drivers, conventional MINI drivers, and GE program participants in Los Angeles, New York

and New Jersey generated themes and language used in a larger nationwide online survey and design game. The third, a nationwide survey of new car buyers, which included subsets of hybrid buyers and recent EV lessees and owners, was used to justify the addition of two GE products for the BMW Active E.

EV customer preference for GE is very strong, with 29% of Active E customers reporting having home solar. One question remains: Is the BMW Group observing an early market phenomenon or a long term trend and opportunity? The vehicle and energy plan design game results illuminate the benefits to an auto OEM of a combined EV and GE product offering. Though conventional vehicle buyers are motivated by different reasons, fuel cost savings, combining EVs and GE increases the overall demand for EVs by 23% in that new car buyer segment.

EVs and GE are transformative and disruptive because the combination represents a disconnection from current liquid transportation fuel systems along with the decoupling of environmental impacts of vehicle mobility from on-board fuel consumption. GE for the home and EV has the potential to lower total household energy costs, which is appealing for new car buyers across all segments.

References

1. Turrentine, T.S., Garas, D., Lentz, A., Woodjack, J.: The UC Davis MINI E Consumer Study. Institute of Transportation Studies, University of California, Davis, Research Report UCD-ITS-RR-11-05 (2011)
2. Yang, C., McCarthy, R.W.: Impacts of plug-in electric vehicle charging. EM Magazine, Air & Waste Management Association (2009)
3. Lutsey, N.P., Sperling, D.: Regulatory Adaptation: Accommodating Electric Vehicles in a Petroleum World. Energy Policy 45, 308–316 (2012)
4. Delucchi, M.A., Lipman, T.E.: Lifetime Cost of Battery, Fuel-Cell, and Plug-in Hybrid Electric Vehicles. In: Pistoia, Gianfranco (eds.) Electric and Hybrid Vehicles: Power Sources, Models, Sustainability, Infrastructure and the Market. Elsevier, Amsterdam (2010)
5. Anair, D.: State of Charge: Electric Vehicles' Global Warming Emissions and Fuel Costs Savings Across the United States. Union of Concerned Scientists, Cambridge (2012)
6. Ferry, M.: California Plug-in Electric Vehicle Owner Survey. California Center for Sustainable Energy (2012)
7. Kurani, K.N., Axsen, J., Caperello, N., Bedir, K., Tyree-Hageman, J.: Consumers, Plug-in Electric Vehicles, and Green Electricity. Presented at Plug-in Electric Vehicles and Clean Energy in California, Sacramento, CA (2012)

SOCIETY: A Social Reading Application to Join Education and Social Network Experience

Elena Guercio[1], Fabio Luciano Mondin[1], Maurizio Belluati[1], and Lucia Longo[2]

[1] Telecom Italia S.p.A. Torino, Italy
{elena.guercio,fabioluciano,maurizio.belluati}@telecomitalia.it
[2] Politecnico di Torino
lucia.longo@polito.it

Abstract. This paper describes a social reading application designed for education and school communities based on electronic book reading experience. Both design and evaluation followed a user centered approach in order to obtain a custom made application called SOCIETY, to be used by community based groups with similar reading interests, running both on IOS and Android System. A special emphasis was given on personal and social aspects in writing and sharing notes with the community groups and the "facebook or twitter" large community to better understand how that can improve the learning process.

Beside showing a high interest on society@school app our trial evidenced the need to separate school community from wider social networks, a need which seems to be important for students. That stresses the importance of privacy management in school applications design, because students seem to be worried about sharing information and mixing school and social networks contexts.

Keywords: Social reading, privacy/respecting, User Experience, School, User Interface, Personal, Social.

1 Introduction

The number of social network users is expected to grow in the next few years. According to a research performed by eMarketer [1] in 2011 there were 1.20 billion people using social networks, number which is expected to grow to 1,85 Billion through 2014. Social networks are making multimedia content sharing very easy and funny, also due to the possibility of exchanging opinions and thoughts together with digital content. On the content side, expenses related to digital content are probably going to grow as well as delivery and access to digital content becomes easier and easier. The 2011 PwC report [2] forecasted consumer spending for entertainment and media industry to surge from $1.4 trillion in 2010 to $1.9 trillion by 2015. Digital content certainly include e-books, and in particular Children and Young Adult's e-books, which according to the American Association of Publisher showed a massive +475.1% increase from 2011 to 2012 [3]. E-book adds a new dimension to traditional books: interactivity. E-Book enables users interaction both with the media and between users, creating and sharing notes and content on top of the e-book content.

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 277–284, 2013.

Market data and technical innovation aspects seems to open interesting opportunities in the education field, and that is the main reason standing behind the choice of developing SOCIETY (SOCIal Ebook communiTY) application: giving students the possibility to add a note reporting their opinion, summary, or link to external content, making the e-book 'live' and dynamic. A note may contain multimedia content (i.e. audio, video, image), extending the original plain text book as a real and new multimedia dimension.

An application to be used by students should be designed in order to be suitable for this scenario, finding a proper trade-off between a "natural" wider use of social networks and privacy concerns. Sharing content, especially in the education environment, implies also the definition of instruments and rules for identifying 'who is sharing what and with whom' and for moderating the shared extra contents in order to avoid misleading or offensive notes.

Moreover the integration with general social networks (i.e. Twitter, Facebook), should be addressed carefully. Giving right to all user profile information just for posting a message and sharing it on user preferred social network, could be enough for reject the authorization to proceed.

2 State of the Art

2.1 Technical View

As it was said previously the e-book is more than a different way to read: it is a new interactive experience. Reading a book on a tablet gives the user the possibility to perform actions on the book, spreading from basic operations such as changing font style or size, to advanced operations such as adding notes and multimedia content to the original book.

The most popular reading applications, Apple iBook and Amazon Kindle, someway give the user such kind of interactivity. The user can underline text or add textual notes and save them on the cloud in order to find them among their different terminals, but it's not possible to interact with people reading the same book except for the possibility to post a note as a new status on the chosen social network (i.e. Twitter or Facebook). Other applications such as Readmill, Inkling and Kobo extends the possibility to share text annotations among users owning the same e-book and discuss by using comments.

Adding text notes to e-books is anyway just a small part of a wider story. Content could be extended by adding multimedia annotations by giving the users the possibility to add audio, photo and video to be enjoyed using the application itself.

On the other hand, sharing multimedia content requires keen user identification in order to avoid privacy and moderation issues. Teachers should be able to delete inappropriate contents or promote the best contributions as new teaching material for the entire class.

Moreover, a social reading solution should be able to manage content protected by Digital Right Management system in order to make it work also with bestsellers and editorial books.

2.2 User Experience View

As it was also explained previously, compared to traditional paper books, eBooks have several advantages [4]. Digital books can give the possibility to overcome some reading diseases by means of customizable displays, page and font layout changes or by combining text reading with auditory stimuli. E-books are also easy to be transported and always available. According to Abram, these features can revitalize the Past, with a sort of culture and thinking restoration, by re-inventing books in a creative way, through video, animations and sounds.

In the article "Mobile phones in the Classroom: If You Can't Beat Them, Join Them'', Scornavacca , Huff And Marshall [5], by collecting data about 1200 students in a large undergraduate class, showed "that both students and instructors can benefit from an additional channel of communication – SMS messages via mobile phones - in the classroom. The lecturer perceived a gain of quality and quantity of feedback from the students. Students indicated that the system was useful - making classes more interesting and interactive".

It was 2002 when Takeda e Suthers [6] tried to create a social reading web tool (Pink), while in 2008 Kim, Farzan and Brusilovsky [7] evaluated an interface in a classroom study, designed to allow the users to identify spatial regions in textbook pages and use them for bookmarking, commenting, and sharing opinions with others. This work stated that "annotation is a natural and beneficial activity for users because it assists them in reading, writing, sense-making, indexing and retrieving". Both works showed how users seem to need online tools to support discussions and learning by making artifacts, references and annotations in a easily way .

Other researchers, such as Kiili et al. [8] tried to identify different profiles for collaborative reading in order to find out different styles in sharing content and found five different profiles: co-constructers, collaborators, blenders, individually oriented readers, and silent readers.

This is the context in which Society try to take its place. Our prototype is based on the idea that reading can be a social and shared experience and annotations help the reader in text understanding, storage and organization and the reading experience will be made more and more interactive and significant on mobile devices such as tablet PC.

3 Application Description

With SOCIETY (Fig.1) users can share their reading experience with other readers or books' authors. For instance in educational environment they can share their learning experience with classmates and teachers by creating and commenting annotations on schoolbooks both from home and from school, experiencing a new and more dynamic way of collaborative teaching and learning. The design of SOCIETY kept the focus on allowing the re-use of well-known interaction processes such as notifications, text selection, to build a new interaction process in App by "recycling" the previous know how and by using it in a simple and quick way. So that users can choose

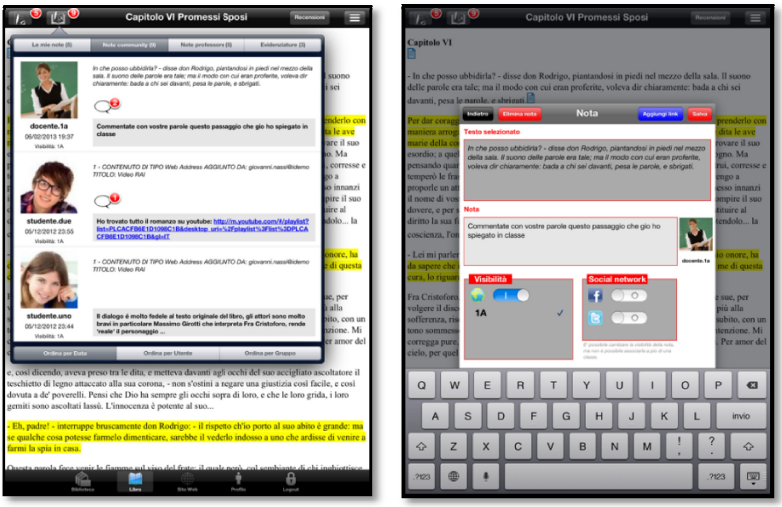


Fig. 1. Application screenshots

the right level of visibility of their annotation, decide if they have to be private, shared only with the author, within a group, or within the whole community and larger social network (i.e. facebook or twitter) too.

SOCIETY was then developed as a prototype App, keeping the user at the center of the design process, and once developed it was tested with final users.

In designing SOCIETY the attention was kept on the actions performed by users using a paper book trying to rebuild this user experience in a simpler, digital way: then, using SOCIETY, the user does not only read e-books, he also underlines text, put notes beside the text, he sometimes places bookmarks, keeps it, lends it, talks about it with his friends.

4 User Requirements and Field Trial

In general e-books make learning easier through collaboration. The students often prefer giving their comments to a text by writing them due to "social" implications of speaking with their teachers and classmates.

In designing Society@school app we had to take into account both data coming from trials and literature. We also involved some teachers and the Director of an high school in Trento, through some open face to face interviews.

The first challenge, was to combine basic requirements (text customization, annotation sharing, access from smartphone and web, etc) with some more specific requirements, such as having separate views for private and public notes was highlight.

The extension of school contents towards large communities (facebook and twitter overall) seemed to be a very promising opportunity as "everyday students usually study and do their homework with Facebook open". Social Network sharing could be

important in order to create a link between the students' school and private life, stimulating them this way to write and share contents about school texts or reading books.

The App was developed both for IOS and Android with the perspective to reach as many device as possible, keeping anyway the same approach to logical design of GUI information, accessibility, navigation, number and types of functionalities and trying to give a common look and feel to both apps.

Once developed, the app was tested by means of a six- months field in Trento, in the same school in which we collected most of the requirements.

The users' sample involved 42 students in two different classes and 3 teachers. The trial started with s short presentation of the App to the students and teachers both to explain the App focal points and usages, and to support the users with the installation on their devices.

Afterwards, we constantly monitored the app usage in different contexts with different teachers, trying to collect as many information as possible from the anonymous log files, but also getting explicit and direct opinions or comments through a specific in-app feedback functionality. Finally we used a poll to get a general opinion about the whole trial and the App's functionalities and performances. In order to stimulate feedbacks a specific web portal section was developed for the description of the development team actions related to users' feedback.

The data collected during the field trial could be classified as:

- spontaneous input and suggestions from the students during the App interaction by email and a specific App software button
- students log files (access number, more clicked functionalities, etc…)
- group interviews (about 10 in 2 classes) with students (Fig. 2)
- 3 face to face interviews with teachers (Fig. 2)
- interviews with students' parents (6)
- final questionnaire

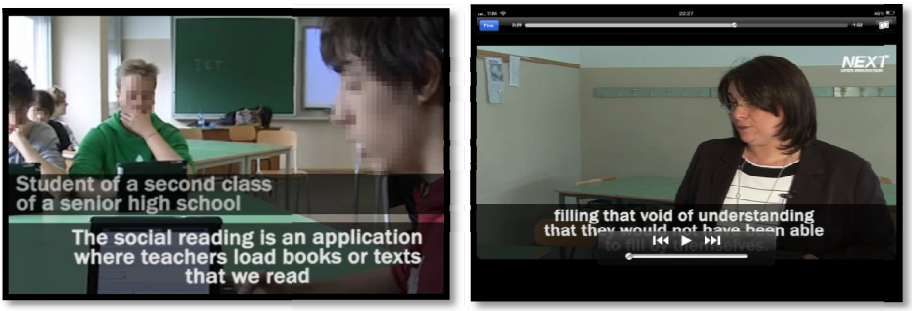


Fig. 2. Face to face interviews with students (on the left) and teachers (on the right)

We used this information to evaluate the usefulness of the App in the school context, but also to optimize and re-design the App and to make it users-tailored. A lot of interesting suggestions were collected, the most important were:

- Notes were not enough: comments over notes were introduced
- explicit highlighting of the notes' cited text.
- Number of notes and comments shown in library section of the app.
- Visibility of notes on class basis.

Unexpectedly "sharing activities to the large communities of Facebook and Twitter" (i.e. not default sharing of notes towards community) from a student perspective, should be kept apart from the friends community making them two worlds to be at least interconnected but not tied.

Students asked for the same social approach and functionalities of their preferred social network, but they adopted a different behavior during the trial. They used to share school data and experience just inside the class and use Facebook and Twitter for extra scholastic experiences: independently from the fact that they received a positive mark from the teacher or even if a note had a lot of comments from the classmates.

We also offered the student a "sign up" functionality via Google, Facebook and Twitter, in order to simplify the registration process, but in nearly every case students weren't interested in using it.

User were concerned about the number and nature of grants and permission to give to the app on Facebook or Twitter. It's important to ask for the minimal and necessary grants in merging the personal and school "environments". Asking access to the entire user profile information just for posting a single message (i.e. user status) seem to be a strong concern for the users and should then be avoided. It's necessary instead to ask for the minimum possible number of grants and explain well why you are asking for it.

To make clear this design choice for the user, we put the default to the direct connection to Facebook and Twitter "off".

In general the survey showed a satisfactory experience (Fig.3) with Society@school: the "vote" of the app was 7.3 on 1 to 10 scale. App and the shared suggestions and optimizations on App proposed both by teachers and pupils were taken into account in the new App version.

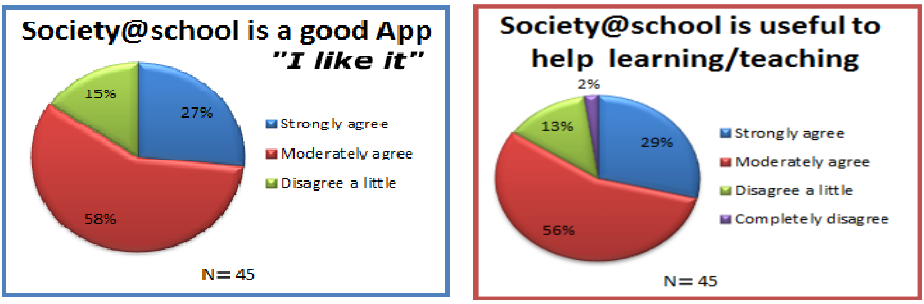


Fig. 3. Users' comments on Society@school App

5 Conclusions and Next Steps

Our trial results show us how high has been the interest on society@school app, as the figure n. 4 shows (Fig. 4).

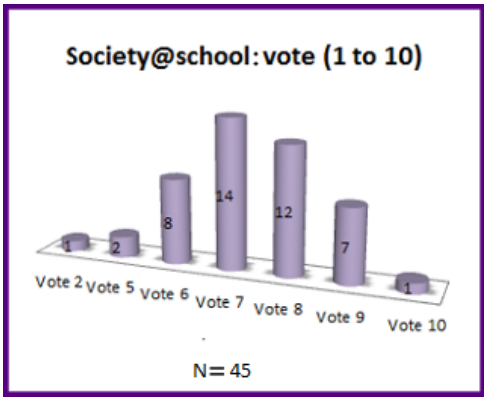


Fig. 4. Final evaluation of society@school app

Over 85% of users have expressed a positive opinion on it, highlighting the added value to the traditional learning. Trialists have also judged the app very user friendly and easy to use. In fact, in the interviewed sample, 82% of them have evaluated easy to read notes, 71% easy to share them, and still 71% easy to insert new one or edit it.

Very important is the "discovery" that, more for students than for teachers, a separation between school community and the wider social networks community is necessary. Thanks to the log files we traced, we have found that only a few percentage of users (about 10-15%) has linked own society@school account to the Facebook or Twitter profiles. Moreover, even among those who have linked society@school profile to Facebook or Twitter, few users have posted their own society@school activities. This finding suggests how important is the privacy management in school applications design, because people, especially teenagers, seem to be very cautious and worried about sharing information and mixing school and social networks contexts.

Our trial has been also very useful to optimize our app, adding new features requested by users such as the opportunity to comment others' notes or to highlight text you are reading. Then about privacy, the "shrewdness" we adopted:

- the new inserted note has to be checked by the student before be shared with others (instead of that the note is just "private")
- then, in a second step, student can decide if sharing or not his note on Facebook and Twitter

The next step of this experimentation is to extend the trial to other primary and secondary schools (Bolzano, Bologna and Caltanissetta are starting at the time of writing), to other classes in the same school in Trento and also to some university courses and community environments.

The new trials should focus also on the Android App (which was designed following the same device principles, but it has not been evaluated on the field yet).

It could also be useful and valuable to insert some specific didactical instruments (cognitive maps, vocal recording or reading, test customized formats, etc…) in order to extend the app with some tools helping dyslexic students both at school or home with the final aim of minimizing the individual difference in reading and learning for the student suffering diseases. In a context like the Italian school in which the number of students suffering DSA problems is increasing, social reading could be a way to compensate some learning diseases. At the time of writing we are collecting requirements in order to extend the app with tools for students suffering DSA.

References

1. http://www.workinghomeguide.com/10763/huge-social-network-users-growth-worldwide-by-2014-facebook-reinforce-dominance
2. http://blog.highbeambusiness.com/2011/08/pwc-report-digital-content-spurs-global-consumer-spending-growth/
3. http://www.publishers.org/press/62/
4. Stephen, A.: P-Books vs. e-Books: Are there education issues? Multimedia and Internet @ Schools (November/ December 2010)
5. Scornavacca, Huff, Marshall: Mobile phones in the Classroom: If You Can't Beat Them, Join Them. Communications of the ACM 52(4) (April 2009)
6. Suthers, T.E.: Online Workspaces for Annotation and Discussion of Documents. In: ICCE International Conference on Computers in Education (2002)
7. Kim, Farzan, Brusilovsky: Social Navigation and Annotation for Electronic Books
8. Kiili, et al.: Working on Understanding During Collaborative Online Reaing. Journal of Literacy Research 44(4), 448–483 (2012)

The Evaluation of a Voting Web Based Application

Linda Harley, Keith Kline, Jerry Ray, Carrie Bell, Andrew Baranak,
Chandler Price, Matthew Hung, and Brad Fain

Georgia Tech Research Institute, Atlanta, Georgia, USA
{linda.harley,keith.kline,jerry.ray,carrie.bell,andrew.baranak,
chandler.price,matthew.hung,brad.fain}@gtri.gatech.edu

Abstract. Ballot layout and the incorporation of assistive technologies into voting systems are plagued with inconsistencies across the United States. The purpose of this study was to evaluate both ballot layout display configuration (information density) and a variety of controllers (e.g., mouse, 2- or 5-button controller) in order to assess performance and preference among voters. Participants were presented with three mock ballots, each with different layouts (scrollable pages, multiple columns or multiple pages per contest). Eye-tracking data and selection time data were recorded and a usability questionnaire was administered after each testing condition. The results of the study found that participants preferred the multiple column display configuration and the use of the mouse. The results from this study will be leveraged to design an iPad Voting Application with appropriate interfaces and controls. This will allow individuals with disabilities the opportunity to vote without requiring the dexterity to use a paper and pencil ballot.

Keywords: vote, web based application, human factors engineering, interface.

1 Introduction

The goal of the *Military Heroes Initiative and Military Voting* project is to better understand and improve voting technology and voting processes that affect recently injured military personnel who have returned from a combat zone with one or more disability. According to statistics compiled by the U.S. Department of Defense, as of February 15, 2013, 50,476 U.S. troops have been wounded during the conflicts in Iraq and Afghanistan, many returning home with a range of disabilities including loss of limbs, impaired vision, and traumatic brain injury [1]. These individuals may have difficulty getting to the polls to vote or even casting their vote via absentee ballot due to their functional and cognitive limitations.

To provide alternative solutions for absentee voting, GTRI developed a web-based voting application test-bed (Voting App), for use with smart technologies such as a smart phone or portable computer. This technology could facilitate obtaining an absentee ballot from an individual who is incapable of visiting the polling station. The purpose of the Voting App is to provide an alternative way of completing the absentee

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 285–294, 2013.

ballot for those individuals who do not have the physical or mental capability to complete a standard paper and pencil absentee ballot. In the Voting App test-bed two variables were considered: (1) information density and (2) controllers.

1.1 Information Density

With an increase in digital reading, from reading on computers at work to leisurely eBook reading, finding an optimal presentation of information is important. To date, research investigating the optimal use of line length, multiple columns, and text justification is inconclusive. Longer line lengths result in faster reading times [2-5], but research suggests medium to short line length may result in greater comprehension [6-7]. Objective results support the use of both single columns of text [4], and multiple short columns [8] but subjective preference seem to be multiple short columns [4], [9-10].

Alternatively information can be displayed on multiple pages or on a scrollable page. Paging refers to navigating between pages where all text is replaced by a new screen full of information. Scrolling involves one page where text is replaced when new text appears on the screen. The objective is to ensure that individuals can move from page to page as efficiently as possible. If designers are unable to decide between paging and scrolling, it is recommended that they provide several short pages rather than one or two long pages that involve significant scrolling. However, with pages that have fast loading times, there may be no reliable difference between scrolling and paging when people are reading for comprehension [11]. The results of this study should help provide enough information to make confident decisions as to the appropriate layout of information.

1.2 Controllers

The most complete collections of hand anthropometry data have been collected by the US Army [12-13]. Studies have shown no significant difference in anthropometric hand measures between people with and without disabilities [14] or due to aging [15]. However, more longitudinal data are needed to fully confirm the aging effects. Where the considerable differences are generally seen is in the areas of strength and motor control, for example, with age there are considerable decreases in strength, dexterity, precision, coordination, joint mobility and sensitivity [16-17].

Push buttons have been shown to have a small space claim and are easy to operate [18]. According to recent occupation injury data, finger injuries are the most prevalent of upper extremity injuries [19]. While performance with controls such as knobs and dials are severely affected, push buttons show no performance difference between disabled and able-bodied users [20]. To better determine differences in button configurations two separate push button configurations were created for this test. The first of these configurations used two buttons, allowing for a *tab* and *select* control (Fig. 1. (A) 2-Button and (B) 5-Button configuration). The other configuration was a five button system, which allowed for *tab*, *shift* + *tab*, *left* (mapped to previous page command of the Voting App), *right* (mapped to the next page command of the Voting App), and *select* (Fig. 1. (A) 2-Button and (B) 5-Button configuration).

1.3 Hypotheses

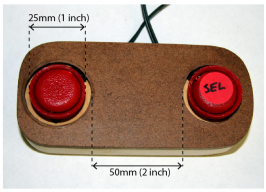
The purpose of this study was to examine information density and controllers for the Voting App.

Aim 1
Investigate the effect of different information density displays for the Voting App.

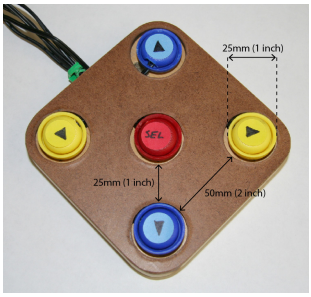
Hypothesis 1.1: The null hypothesis is that there is no significant difference between scrollable (*scroll*), multiple column (*columns*) or multiple page (*pages*) display configurations.

Aim 2
Investigate the effect of different controllers for the Voting App.

Hypothesis 2.1: The null hypothesis is that there is no significant difference between the *mouse*, *2-button* and *5-button* controls.



A



B

Fig. 1. (A) 2-Button and (B) 5-Button configurations

2 Method

2.1 Participants

Eighteen (18) participants (26±13 years of age, 12 males and 6 females) volunteered for this study. Participants were 18 years of age or older and considered to be healthy. The Georgia Institute of Technology Internal Review Board approved the study.

2.2 Experimental Setup

Participants were asked to sit comfortably behind a desktop computer (Fig. 2. Participant using the 2-button controller to navigate the Voting App at the eye-tracking station.). The SmartEye system was used to track the participant's eye movements in order to determine areas of fixation on the screen. Three SmartEye cameras (Basler acA640-100gm cameras with 8mm lenses), each with two IR flasher devices, were used to track eye movement at a sampling rate of 60Hz. Three controllers were used

in this study: (1) *mouse*, (2) *2-button* controller (Fig. 1. (A) 2-Button and (B) 5-Button configuration) and (3) *5-button* controller (Fig. 1. (A) 2-Button and (B) 5-Button configuration). The buttons (Enabling Devices Compact Switch #745) were connected to a switch interface device (X-keys XSI-38-US). The controllers interfaced the 3.5 mm switch plugs to the PC via a USB port, and facilitated programming the switches with the desired keyboard inputs. These buttons were mapped to standard keyboard keys, which were then encoded in the Voting App.

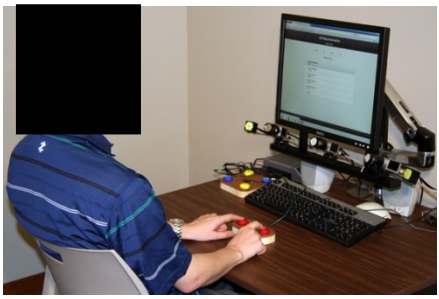


Fig. 2. Participant using the 2-button controller to navigate the Voting App at the eye-tracking station

2.3 Experimental Groups

Participants were pseudo-randomly assigned to one of three groups so that age and gender were balanced between groups. This study used a 3x3 factorial design, with the different controllers as a between-subjects group variable and the different display configurations as a within-subject variable. Each participant, assigned to only one of the three controllers, was asked to complete the mock ballot under each of the three display configuration conditions. The order in which the display configurations were tested was counterbalanced across participants by Latin square to eliminate order effects.

2.4 Task Protocol

After consent had been obtained, participants were asked to sit at the SmartEye computer while the cameras were calibrated to their individual anthropometric characteristics. After calibration, the participants were instructed on how to use the controller and presented with one of the ballot configurations. Participants were instructed to experiment with the ballot and take their time. After completion of a ballot, participants completed the System Usability Scale questionnaire. Then participants were given a few minutes to rest, while the SmartEye system saved the data. The process was then repeated using a different ballot configuration. At the completion of the study, participants were asked several interview questions to determine any additional difficulties they may have encountered.

2.5 Data Analysis

Time Dependent and User Selection Variables

The Voting App contained a built-in function that logged each keystroke and selection made by participants along with the time stamp. The following variables were calculated:

- Ballot Duration: The total time it took the participant to complete the ballot. Time started when the participant clicked on the "Start" button and stopped when the participant clicked the "Submit" button.
- Mean Click Time: Click time is defined as the time between selections made by the participant. Click time was averaged across the entire ballot to calculate Mean Click Time.
- Overvote: An overvote occurred when the participant attempted to vote for more participants than they were allowed to for that particular contest. The number of overvote's that occurred for the entire ballot was tabulated.
- Undervote: An undervote occurred when the participant attempted to vote for less candidates than they were allowed to for that particular contest. The number of undervotes' that occurred for the entire ballot was tabulated.
- Back Button: The number of times the "Back" button was selected (Fig. 3. Screenshot of ballot layout).
- Help Button: The number of times the "Help" button was selected (Fig. 3. Screenshot of ballot layout).
- Review Button: number of times the "Review" button was selected. This button allowed participants to review their entire ballot (Fig. 3. Screenshot of ballot layout).
- Next Button: The number of times the "Next" button was selected. This button allowed participants to go to the next page (Fig. 3. Screenshot of ballot layout).

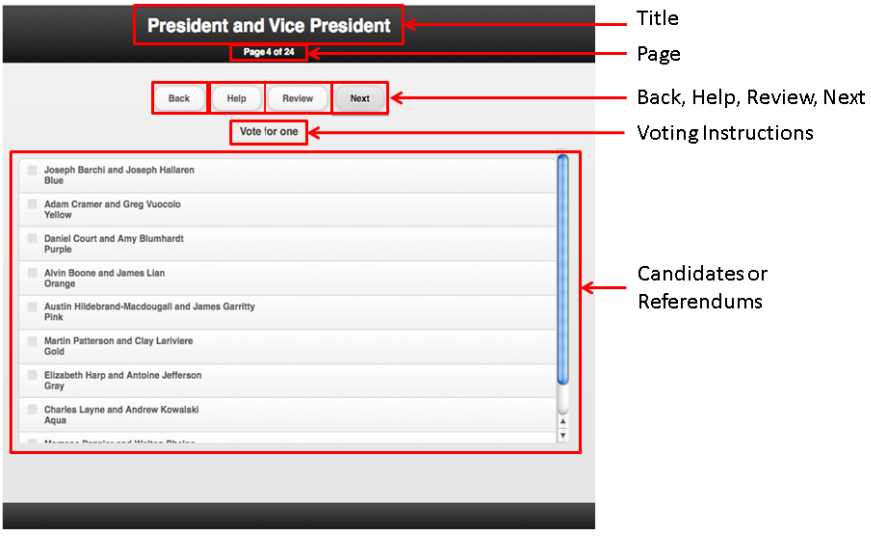


Fig. 3. Screenshot of ballot layout

Eye Tracker Dependent Variables
The eye tracker system was used to determine how much time participants spent viewing several regions of interest on the Voting App. The following variables were recorded and analyzed using the SmartEye software. The look zones for the study began recording data when the participant was presented with the first page and stopped recording once the user submitted the ballot. The time spent viewing each of the following regions of interest were monitored (Fig. 3. Screenshot of ballot layout): (1) Candidates and referendums, (2) Title, (3) Voting Instructions, (4) Page, (5) Back, (6) Help, (7) Review, (8) Next.

3 Results

3.1 Time Dependent and User Selection Variables

A two-way ANOVA was run with display configuration (*pages*, *column* and *scroll*) and controls (*5-button*, *2-button* and *mouse*) being the independent variables. The statistical analysis found no significance between the independent variables for the ballot duration, mean click time, number of undervotes, or number of times that the "Help" button was selected ($p>.05$ for all comparisons). However, a significantly greater number of overvotes occurred for the *pages* display configuration when compared to both the *column* and *scroll* display configurations, $F(2,30) = 8.73$, $p = 0.001$. This is supported by the fact that the number of times the "Back" and "Next" buttons were selected was significantly greater for the *pages* display configuration when compared to that of the *column* and *scroll* configurations, $F(2,30) = 80.78$, $p<0.001$. For the controllers, the only significance among conditions was that the "Review" button was selected more frequently for the *2-button* controller than the *5-button* and *mouse* controllers, $F(2,15) = 5.98$, $p = 0.012$. There were no significant interactions between display configurations and controls.

When considering the difference between display configurations, these results suggest that many more errors occurred for the *pages* display configuration than the *columns* and *scroll* configurations. This resulted in a greater amount of overvotes occurring, meaning an increased navigation between *pages* as far as the number of times that the "Next" and "Back" buttons were selected. One possible explanation may be that cognitive load was high for the *pages* display configuration, since participants were using working memory to recollect how many candidates they had selected for the contest. When evaluating the difference between controllers, the only result of note was that participants in the *2-button* condition more frequently selected the "Review" button. This strategy may have assisted participants in avoiding unnecessary navigation between the pages, as the *2-button* controller only allowed for forward linear advancement and selection. Although not significant, participants using the *mouse* had the shortest ballot duration and mean click times. These results indicate that the most optimal controller was the *mouse* and that the *pages* display configuration was not an optimal display configuration. There were no significant differences between *scroll* and *column* display configuration.

3.2 Eye Tracker Dependent Variables

A two-way ANOVA was run with display configuration (*pages*, *column* and *scroll*) and controls (*5-button*, *2-button* and *mouse*) as the independent variables, and time spent viewing regions of interest on the voting app as the dependent variable. The statistical analysis found no significance difference between the independent variables for the time spent viewing the "Candidates and referendum", "Voting instructions", "Page numbers", "Back", "Help", and "Next" buttons. However, participants spent a significantly greater amount of time viewing the "Title" for the *pages* display configuration when compared to the *column* and *scroll* display configurations, $F(2,20) = 10.99$, $p < 0.01$. In addition, results showed that participants spent more time viewing the "Title" for the *5-button* controls than the *2-button* controls. In fact, the time participants viewed the "Title" for the *2-button* controls was small and negligible when compared to the other conditions. This may be due to the linear nature of the navigation.

These results support that of the previous section, in that participants in the *pages* display configuration had to pay much closer attention to the title of the page to understand which contest they were on.

3.3 Subjective Rating Results

System Usability Scale

A 2-way analysis of variance was conducted on the System Usability Scale data, with display configuration as a within-subjects variable and controls as a between-subjects variable. The data are shown in Fig. 4. Subjective ease of use ratings obtained with

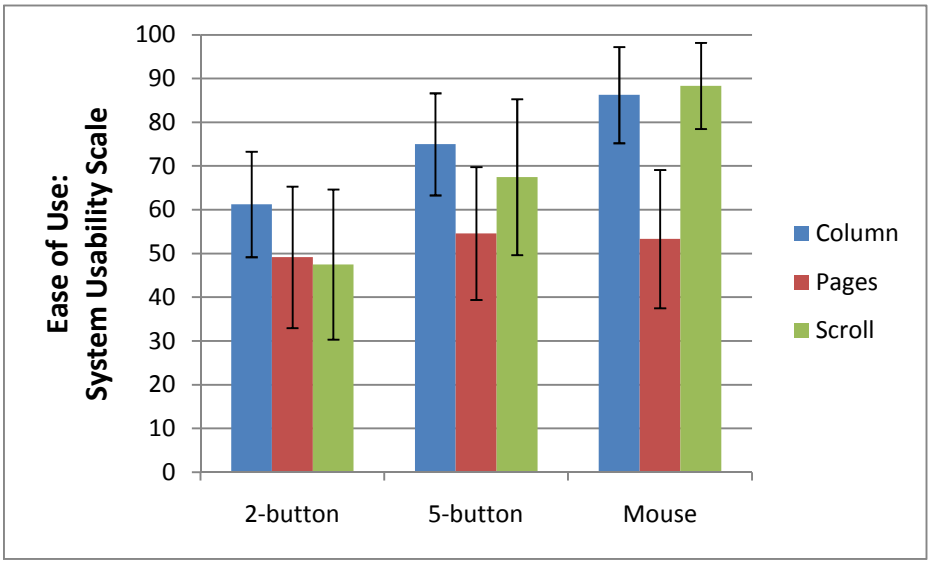


Fig. 4. Subjective ease of use ratings obtained with the System Usability Scale. Error bars represent 95% confidence intervals.

the System Usability Scale. Error bars represent 95% confidence intervals The effect of display configuration was significant, $F(2,30) = 7.99$, $p = 0.002$. The effect of controls was not significant, $F(2,15) = 3.15$, $p = 0.07$. The interaction between controls and display configuration was not significant $F(4,30) = 1.82$, $p = 0.09$.

Post-hoc paired t-tests were conducted on the display configuration conditions. The critical p-value was adjusted to 0.0167 for multiple comparisons. Ratings for the *column* condition were higher than those in the *pages* condition, $t(17) = 4.95$, $p<0.01$, but ratings for the *column* condition were not significantly higher than those in the *scroll* condition, $t(17) = 1.02$, $p = 0.32$. Ratings for the *scroll* condition were not significantly higher than for the *pages* condition, $t(17) = 2.30$, $p = 0.03$.

4 Discussion

The purpose of this research was to examine different information density display configurations and different controllers for the Voting Application. The first null hypothesis was that there would be no significant difference between the *scroll*, *columns* or *pages* display configurations. The results of this study suggested that there was no significant differences between display configurations when considering the length of time it took participants to complete the ballot or the mean time between selections. However, the number of overvotes that occurred during the *pages* display configuration was significantly higher than the *columns* or *scroll* display configurations. Participants also selected the "Back" and "Next" navigation buttons a significantly greater amount of times for the *pages* display configuration when compared to the other conditions. This implies that perhaps the cognitive load on recalling vote selections as well as understanding what contest users were viewing was too complex. This is supported by the fact that participants spent an increased amount of time looking at the "Title" of the contests for the *pages* display configuration than the other conditions. Moreover, results from the System Usability Scale showed that participants had difficulty with the *pages* display configuration. Participants reported that the *pages* display configuration was by far the most difficult, confusing and memory intensive. There were no significant differences in performance of participants between the *scroll* and *column* display configurations. Therefore, the research concludes that the *pages* display configuration is the least optimal solution for presenting information on a ballot, and that either *scroll* or *column* display configurations should be used.

The second null hypothesis was that there would be no significant difference between the *mouse*, *2-button*, and *5-button* controls. Participants utilized the "Review" button more often in the *2-button* condition when compared to the *5-button* or *mouse* conditions. This approach may have been useful in avoiding cycling through all the contests to select only those contests in which they were interested. Recall that the *2-button* controls only allowed for one-directional (tab-forward) movement throughout a page. Furthermore, participants spent little time looking at the "Title" for the *2-button* control (492±225 milliseconds). If participants navigated to the contest from the review page, they would not need to read the title, as they would already know what contest they were going to vote on. Participants reported that the *2-button*

control was cumbersome. The *mouse* control was the easiest to use, which is likely due to the fact that participants are already familiar with this technology. These results lead us to conclude that the *2-button* control is the least optimal solution, while the *5-button* control is mediocre and the *mouse* is preferred. A more optimal button solution may be 3-button controller, since it will allow forward and backward navigation as well as a selection button. This would be less demanding than the *5-button* controller and more flexible than *2-button* controller.

One of the main goals of the Voting App test-bed is to utilize it to evaluate a variety of long-standing issues being discussed in the voting world, such as the use of plain language in instructional content, font types, cognitive load, symbolism, and so forth. In addition to the Voting App test-bed, research staff is also in the process of designing an iPad hard case with a 3-button interface to be used in testing.

References

1. U.S. Casualty Status for OI, OND, and OEF, `http://www.defense.gov/news/casualty.pdf`
2. Duchnicky, R.L., Kolers, P.A.: Readability of text scrolled on visual display terminals as a function of window size. Human Factors 25, 683–692 (1983)
3. Dyson, M.C., Kipping, G.J.: The legibility of screen formats: Are three columns better than one? Computers and Graphics 21(6), 703–712 (1997)
4. Dyson, M.C., Kipping, G.J.: The effects of line length and method of movement on patterns of reading from screen. Visible Language 32, 150–181 (1998)
5. Dyson, M.C., Haselgrove, M.: The influence of reading speed and line length on the effectiveness of reading from screen. Int. J. Human-Computer Studies 54, 585–612 (2001)
6. De Brujin, D., De Mul, S., Van oostendorp, H.: The influence of screen size and text layout on the study of text. Behavior and Information Technology 11, 71–78 (1992)
7. Chaparro, B., Baker, J.R., Shaikh, A.D., Hull, S., Brady, L.: Reading online text: A comparison four whitespace layouts. Usability News 6(2) (2004)
8. Lam, K., Lam, Y., Liu, J., Shin, U.G.: Reading comprehension and rate: one column versus three columns. Dept of Computer Science Class Project, Univ. of Maryland, College Park (2000)
9. Bouma, H.: Visual reading processes and the quality of text displays. In: Grandjean, E., Vigliani, E. (eds.) Ergonomic Aspects of Visual Display Terminals. Taylor & Francis, London (1980)
10. Andrevey, V., Martynov, A.: Effects of splitting text into multiple columns. Dept of Computer Science class Project, Unv. Of Maryland, College Park (2000)
11. U.S. Department of Health and Human Services: The research based web design and usability guidelines (2006)
12. White, R.M.: Comparative anthropometry of the hand. U.S. Army Natick Research & Development Laboratories (1980)
13. Grenier, T.M.: Hand anthropometry of U.S. Army Personnel. U.S. Army Natic Research, Development, and Engineering Center (1991)
14. Feeny, R.: Specific anthropometric and strength data of people with dexterity disability. In: Consumer and Competition Policy Directorate: Department of Trade and Industry (2002)
15. Molenbroek, J.F.M.: Anthropometry of elderly people in the Netherlands; research and applications. Applied Ergonomics 18(3), 187–199 (1987)

16. Daams, B.J.: Human force exertion in user-product interaction. Doctoral Thesis, TU Delft, Amsterdam (1994)
17. Haigh, R.: The ageing process: a challenge for design. Applied Ergonomics 24(1), 9–14 (1993)
18. Van Cott, H.P., Kinkade, R.G.: Human engineering guide to equipment design. United States Dept. of Defense, Joint Service Steering Committee (1972)
19. Bureau of Labor Statistics: Nonfatal occupational injuries and illnesses requiring days away from work. 2010 US Department of Labor (2011)
20. Pennathur, A., Mital, A., Contreras, L.R.: Performance reduction in finger amputees when reaching and operating common control devices: A pilot experimental investigation using a simulated finger disability. Journal of Occupational Rehabilitation 11(4), 281–290 (2001)

A Refuge Location Prediction System for When a Tsunami Has Occurred

Akhiro Kawabe, Tomoko Izumi, and Yoshio Nakatani

The Graduate School of Science and Engineering, Ritsumeikan University,
Kusatsu city, Shiga, 525-8577 Japan
is013087@ed.ritsumei.ac.jp,
{izumi-t,nakatani}@is.ritsumei.ac.jp
http://www.sc.ics.ritsumei.ac.jp/

Abstract. During the 2011 Tohoku Earthquake and Tsunami, DMATs (Disaster Medical Assistance Teams) could not rescue victims efficiently with accurate location data, because the local governments had lost refuge location data and resident registers due to damage caused by the tsunami. In this paper, to support DMATs, a refuge prediction system based on the characteristics of disaster, landscape, and victims' psychology is proposed, which can function even if local governments lose information about victims and refuge locations. As an example, this system deals with tsunami. We demonstrate the effectiveness of this system by comparing the data of the 2011 Tohoku Earthquake and Tsunami and our prediction system.

Keywords: Disaster Medical Assistance Team, earthquake, local government.

1 Introduction

All over the world, suffering is caused by various kinds of natural disaster, including typhoons, floods, volcanic eruptions, and earthquakes. Japan is particularly affected by earthquakes, as the country is surrounded by four tectonic plates.According to recent statistics, magnitude 2.0 earthquakes occur in Japan about 10,000 times a year.For that reason, Japan is a country of frequent earthquakes and many earthquakes happen every year[1].

The 2011 Tohoku Earthquake and Tsunami occurred on March 11, 2011, killing more than 15 thousand people. This earthquake'smagnitude was9.0, which ranks fourth highest in global statistics of the past 100 years.

In the 2011 Tohoku Earthquake and Tsunami, many hospitals and medical institutionswere damaged by the tsunami, causing shortages of doctors. To solve this issue, DMATs (Disaster Medical Assistance Teams) supported the disaster area. In spite of the DMATs' early dispatch to the disaster area, they were not able to rescue with accurate location data, due to the fact thattown office administration stopped becausemany coastal town offices suffered from the tsunami. There were no means of contacting victims, as the town offices had lost records of the sufferers' addresses[2].

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 295–300, 2013.

It is predicted that the probability of a Nankai or Tonankai Earthquake occurring in Japan within 30 years is over 60 percent [3].In future, when earthquakesoccur, it is vital to know victims' addresses, conditions and ages in order to make full use of DMATs' abilities.

In this study, we proposea refuge predictionsystem for supporting disaster medicine from the viewpoint of disaster characteristics, regional characteristics and psychological characteristics, which can be used even when local governments are not able to acquire information on victims' addresses and conditions.

2 Related Work

During a disaster, it is important to exchange disaster information. In the 2011 Tohoku Earthquake and Tsunami, DMATswere not able to acquire sufferers' addresses or refuge location information, as the local governments that should have performed this task were too damagedby the tsunami. In order to solve these problems, DMATs must establish who is in need of rescue, where they are, and acquire refuge location data. If information cannot be obtained from sufferers, we need to predict refuge locations in some way. In the 2011 Tohoku Earthquake and Tsunami, such support was not provided.

Over the past few decades, a considerable number of studies have been conducted on predicting refuge locations. In this section, we briefly introduce such research.

In 2009, Asakawaet al. proposed a system that shares the location information of the user by displaying it on a map [4]. This system attaches these data to Google Maps using a mobile phone camera and GPS function. By sharing the data as map information, the system links between Google Maps and the physical world.Thus, Asakawa et al,'s system proposes an environment in which information can be exchanged as a reality.

In 2007, Tanida and Daito proposed a method for specifying victims' locations using IC tags when the telephone network fails in the damaged area, and demonstrated the system's effectiveness by simulation [5]. In their system, a helicopter scatters active type IC tags on the ground, and the system specifiesvictims' locations by trilateration with three fixed IC tags and the location-unknown subject's IC tag.

In 2011, Google constructed the "Google Crisis Response System"in response to the Tohoku Earthquake and Tsunami [6]this system displays refuge locations with markers in Google Maps, and displays the number of evacuees by marker color. In addition, this system can output refuge location data for all forty-seven Japanese prefectures in CSV andKML formats.

In 2011, Iizukaet al. proposed a systemthat collects disaster situation information and showsit on a map when institutions such as universitiesareaffected by disaster[7].

It is thus clear from the above research that it is effective to show disaster information on maps. However, sufferers were forced to take refuge in unexpected places by the Tohoku Earthquake's large tsunami, andrescue attempts by DMATs were hindered because the tsunami destroyed many designated refuge places. Therefore, a system is required that can predict refuge locations in advance, and give this data to DMATs, thus supporting disaster medicine.

3 System Construction

We proposed “The refuge place forecast system for supporting disaster medicine “as previous study stage.

In this study, we proposeda refuge predictionsystem for supporting disaster medicine from the viewpoint of disaster characteristics, regional characteristics and psychological characteristics, which can be used even when local governments are not able to acquire information on victims’ addresses and conditions from huge tsunami [8]. A previous study stage of system output is shown in Figure 1.

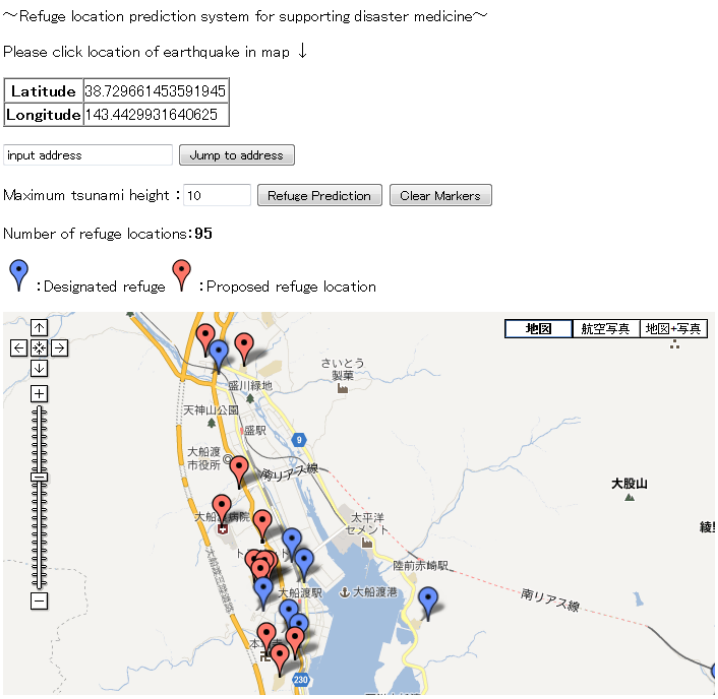


Fig. 1. Output example

In this previousstudy, the user inputs expected maximum tsunami height, refuge place are not displayed if tsunami height is higher than the refuge location’s altitude, sothis system shows refuge location names and the number of evacuees by comparison between regional population and refuge locationpredictionresultif a marker is clicked, as shown in Figure 2.

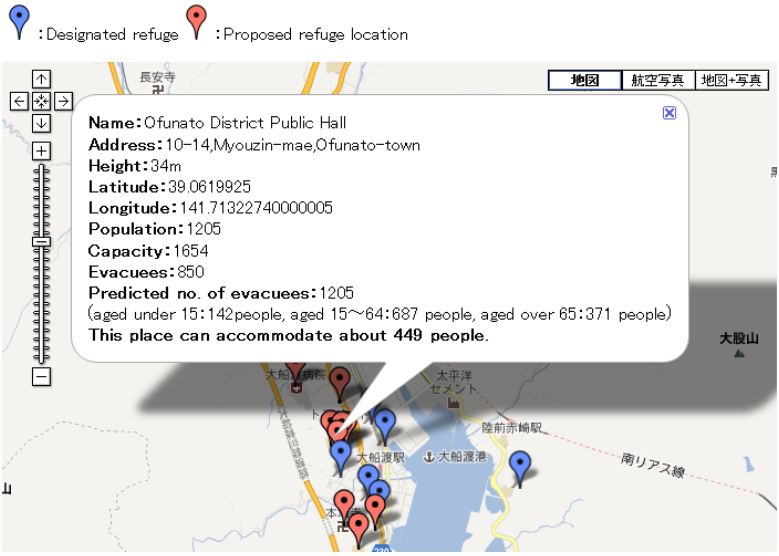


Fig. 2. Output example: marker detail

According to the results of the system's evaluation by Konan Fire Department and Kyoto City Fire Department, it would be useful in future to predict the number of victims, their locations and ages when DMATs are dispatched for early rescue in disaster medicine, however refuge location prediction isn't precise about absence of sufferers' addresses or refuge location information area because this system treats past disaster sufferers' addresses or refuge location information, and this system treats only national census population and age ratio for refuge location prediction.Therefore, in this paper try to accurate predict about the area hadn't occurred disaster.

In this paper, we propose system based on previous study problem not only accurate predict about the area hadn't occurred disaster, but also assume aboutNankai,Tonankaiearthquake will occur. After, we aim to build a platform in which local residents can mutually register the information. For example, local residents could register refuge locations and places where in the past ground liquefaction had occurred, and we propose that system predict refuge place based on registered information when earthquake occur. In Tohoku Earthquake, many evacuees took refuge unexpected place. For example, a skating rink, a private house, a shrine, consequently, DMATs needs get the information that local residents input unexpected place [9]. Also, local residents again realized each other to register unexpected place. For example, local residents think "We didn't know there places", "We can take refuge this place". This system can add choice of refuge location based on refuge location information if tsunami occurred.

4 System Proposal

4.1 About Prototype System

About prototype system register information, local residents input information by smart phone, and prototype system manage database. Local residents register information with text. Register information are time, name, capacity, also this prototype system can upload photo of before disaster refuge place for rescuing if DMATs have visited disaster area first time. This system shows stamp variety of attribute to register refuge place, and can display refuge locations information in map by these stamp.An image of prototype system is shown in Figure 3.

Fig.3. An image of prototype system

4.2 System Flow

In this research is composed of local residents, disaster countermeasures office to pass date. The process of the system consists of the following web application tasks.

① Local residents inputs information about refuge location, and register information for database.
② This system abstracts refuge location information (capacity, evacuees)from database.

③ This system compares abstracted data and predictive refuge location of previous study, and modifiesits data with real-world circumstances so that DMATs can get accurate information.
④ This system sends modified data to disaster countermeasures office.
⑤ DMATs start rescuing based on modified data.

It is thus that this system can support DMAT because input data of local residents and disaster countermeasures office cooperate. To support DMAT is Disaster medicines make revitalize, finally be able to rescue.

5 Endnotes

In this paper, we proposed a system that predicts refuge locations in times of disaster so that DMATs can rescue effectively based on this information. In large-scale natural disasters such as the 2011 Tohoku Earthquake and Tsunami, DMATs were unable toobtain victims' locations and refuge location information because local governments that held such information had suffered extensive damage from the tsunami. The purpose of this research is to support disaster medicine.

According to problem of previous study,we propose system based on previous study problem not only accurate predict about the area hadn't occurred disaster, but also assume about Nankai,Tonankai earthquake will occur. In future, we would like to strengthen the alliance betweenlocal residents, disaster countermeasures office to feedback based on information of local residents input.

References

1. Okada, Y.: Japanese earthquake map. Tokyo Shoseki, 7-9 (2004)
2. Shimbun, Y.: Special compact editionone month after the Tohoku-Pacific Ocean Earthquake. Yomiuri Shimbun, 64–65 (2011)
3. The Headquarters for Earthquake Research Promotion, http://www.jishin.go.jp/main/p_hyoka02L.htm (accessed August 27, 2012)
4. Asakawa, K., Hirano, K., Tsukada, T., Hayashi, Y., In, Y., Omiya, Y., Hamai, T., Murakami, H.: Collection and presentation of a system for safety information in disasters. ITE Technical Report 33(11), 123–126 (2009)
5. Noriyuki, T., Masatora, D.: Agent Based Modeling for Finding Victims with IC tag
6. Google: Google Crisis Response:A Google.org project, http://shelter-info.appspot.com/maps (accessed August 28, 2012)
7. Iizuka, K., Iizuka, Y., Yoshida, K.: A Real-time Disaster Situation Mapping System for University Campuses. In: Ozok, A.A., Zaphiris, P. (eds.) OCSC 2011. LNCS, vol. 6778, pp. 40–49. Springer, Heidelberg (2011)
8. Kawabe, A., Nakatani, Y.: The refuge place forecast system for supporting disaster. The 74th National Convention of IPSJ 2ZG-3, 4-791–4-792 (2012)
9. Disaster information portal for Iwate, http://www.pref.iwate.jp/~bousai/index.html (accessed December 6, 2012)

Further Benefit of a Kind of Inconvenience for Social Information Systems

Hiroshi Kawakami

Kyoto University, Yoshida Honmachi, Sakyo, Kyoto, 606-8501, Japan
kawakami@i.kyoto-u.ac.jp
http://www.symlab.sys.i.kyoto-u.ac.jp/

Abstract. This paper discusses on social information systems from the viewpoint of Fuben-eki that stands for *FUrther BENEfit of a King of Inconvenience*. Analyzing several examples with benefit of inconvenience elucidates the relation between inconvenience and their benefits that in turn are organized into design guidelines. The guidelines applied to several design fields including social information system design. A few examples of such design are presented in this paper. One of the examples accumulates and shares information on literature, and the others are for establishing local communities.

Keywords: benefit of inconvenience, social information system, system design theory.

1 Introduction

It is well-known that such convenient systems as automated ones do not always guarantee the high performance of human-machine systems. On the contrary, the ability of auto-adaptation of machine is sometimes harmful for human operators. It spoils human ability of skill acquisition and of maintaining their motivation. The result of the spoilage sometimes leads a seedbed of serious problems for safety engineering. FUBEN-EKI [1] can be seen as one of the approach to the solution of those problems.

Not only for safety engineering but also for social information systems, benefit of inconvenience provides several notions. Generally, inconvenient systems require users labor, skill and consumption of cognitive resources. Turning the viewpoint from *requirement* to *allowance*, inconvenience allow users to interact systems with time/effort, and to demonstrate/improve their skill. Such allowance is one of the issue of benefits of inconvenience. Among wide range of the relevant field of FUBEN-EKI, this paper focuses on social information systems. Social information is, as defined in the research field of fundamental informatics [2], exchanged and accumulated through organizing social life. To get leverage out of social information, support systems for accessing such information are insufficient. They may be superficially convenient but spoil benefits of inconvenience.

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 301–306, 2013.

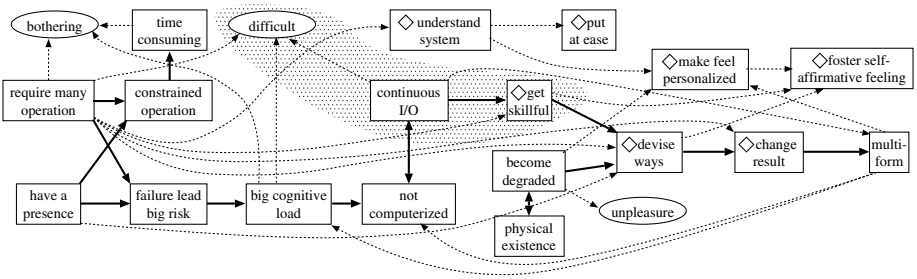


Fig. 1. Relations among inconvenience and benefits

2 Brief Introduction of Benefit of Inconvenience

The procedure for incorporating benefits of inconvenience into systems design consists of

1. defining relevant terms, e.g., benefit and convenience,
2. analyzing, in terms of the above definitions, existing examples that provide users with benefits of inconvenience,
3. generalizing the relationships between inconveniences and benefits that are derived from above analyses and are organized to design guidelines, and
4. applying the guidelines to target systems.

The same procedure is applied in the case of designing social information system.

Naive Definition of Relevant Terms: For common understanding of relevant terms, e.g., convenience, inconvenience, and benefit, we searched for sentences including the term "convenience" in more than 200 web pages. Naively following the results of the web search, this paper defines the terms as follows:

$$\begin{aligned} \text{convenient} \equiv\ & \text{saving labor to attain a specific task} \\ & \text{labor} \supset \text{time consuming} \\ & \text{labor} \supset \text{special skills} \supset \text{mental load.} \end{aligned}$$

Relation between Inconveniences and Benefits: Following the above mentioned definitions, we examined several examples with benefits of inconvenience, and yielded relationships among inconveniences and benefits. Figure 1 overviews a part of such relationships.

In the figure, rectangular nodes denote normal matter, circular nodes denote inconvenient matter, each thick arcs denotes implicational relation, each dashed arcs denotes contribution relation, and $\Diamond$ is an operator of Modal Logic [3]. In the alethic mode, $\Diamond p$ means that the proposition p is possible. In the deontic mode, $\Diamond p$ means that p is permitted. In either mode, even if the value of p varies subjectively depending on humans attitude, the value of $\Diamond p$ is objectively determined.

Utilizing Benefit of Inconvenience: Tracing nodes and arcs in Figure 1 suggests several ways of getting benefit of inconvenience. For example, the hatching part of the figure shows that there are inconvenient tools or methods that are superficially difficult to operate but allow users to get skillful to operate by utilizing continuous I/O.

The nodes in the figure are categorized into three types, i.e., benefits, guides, and strategies. They are organized by means-ends relations. The benefits are

- fostering affirmative feeling,
- providing motivation to tasks,
- personalization [6],
- putting users at their ease.

As shown in Fig.1, we also get some abstract guide for realizing these benefits as follows:

- allowing understand systems (visibility),
- allowing exploration (enhance awareness),
- allowing creative contribution to tasks and unsaturated proficiency.

Furthermore, the strategies for implementing such guidelines have been observed in many examples, e.g.,

- utilize continuity and diversity of physical quantities,
- give a physical presence.

3 Psychological Aspect of Benefit of Inconvenience

Benefits of inconvenience are dependent on individuals. Therefore, psychological notions are beneficial for discussing such benefits.

Living with Complexity: A cognitive psychologist D. A. Norman, known as the presenter of POET [4], declares that adaptation of tools is nonsense and puts emphasis on adaptation of users [5]. He shares the same standpoint with this paper that values such characteristic of tools or methods as "do not disturb users adaptation." He also presented emotional design [6] with the notion of personalization that is one of the benefit observed in inconvenient tools and methods. However, his recent book [7] claims that complexity of operation is inevitable with the modern complex system, and if the system is simple to understand, user should compromise with complexity. On the other hand, this paper regards complexity as one of the key issues to implement benefits of inconvenient. Complexity is not evil but the value of being hard to understand should be carefully examined and feed off the process of systems design.

Ecological Psychology: Inconvenience does not always provide users with benefits. The characteristic of tools or methods, which are generally said to be inconvenient, enables users to aware the possibility of the behavior to get some benefits. In the research field of ecological psychology, affordance is one of the important idea. Affordance is the possibility of behavior, and tunability to affordance differs from one individual to others. Only tuned individual can be afforded to some behavior but not every tuned one behaves according to affordance. The essential quality of benefit of inconvenience is same as affordance. The benefit p is objectively declared $\Diamond p$ that means "p is possible" or "p is permitted." While affordance is directly tuned by animals and affords direct actions, benefit of inconvenience attracts users to relatively long term behaviors, e.g., getting skillful, and devising ways of usage.

Risk Homeostasis Theory: In the research field of human factors, risk homeostasis theory [8] argues that passive restraint devices do not always guarantee the safety because users overly rely on the devices and tend to compensate risks. In other words, convenient automatic safety systems encourage users risky behaviors. On the other hand, inconvenient systems foster users safety behaviors and skills to operate the systems safely.

4 Applying Fuben-eki to Social Information Systems

Applying design guidelines derived from the relationships among benefits and inconvenience, novel social information systems are yielded. Their performance are now under examination in several research fields. This section introduces a few examples of the systems.

Biblio Battle: For accumulating and sharing information on literature, book-review is one of the efficient method and convenient information systems allow us to submit and consult reviews anywhere, anytime and anybody. Biblio Battle [9] is an inconvenient review system that requires participants to physically assemble on a certain time and place. Each participant presents his/her review on his/her favorite book, then participants take a vote on reviews. The inconvenience, i.e., physical assembling provides high quality of review and generates a community via review battles. Biblio battle, proposed by a researcher of Fuben-eki, has grown to become a nationwide event hosted by the Tokyo metropolitan government, and got the award "library of the year 2012" in Japan.

The simple rule of the battle provides high visibility of the battle system and the inconvenience allows to explore strategies to win the battle and to aware the other participants' personalities.

Neko Media: Communities in cyberspace are convenient since getting in and out are relatively easy than physical communities and interaction with community members is not require so much responsibilities. Neko Media [10] is a inconvenient system for establishing a community. Participants are required to be a

friend of a certain stray cat (Neko in Japanese) for taking part of the community because URL of the community site is only known by the QR code attached to the collar of the cat. The difficulty to take part in the community enhances close communications in the community.

Media Biotope: Mass media has enormous influence and is an accomplice of globalization. Media Biotope [11] is an idea to structure media inspired by biotope of ecological system. Based on this idea, Fuben-eki media systems are schemed [12] to localize communication media in a appropriate size of communities. Mediators are potted plants on a footbridge, sign boards on local bus, etc., that promote communications in local societies.

5 Conclusion

It is well-known that not every inconvenience can provide benefits. Even more, not every benefits of inconvenience is acceptable for Fuben-eki. It is compromising that we live with laborious password for security. It is ignominious that complex payment structure makes money. Fuben-eki eliminates such kind of compromises and ignominious situations. Inconveniences should be introduced for only getting benefits as shown in section 2. A promising way to get such benefits is following guidelines as shown in the same section.

Some convenient social information systems are harmful. Job matching sites accumulate recruiting information and facilitate application process. It is convenient because either applicants and recruiters can save many steps. The trouble is that either applicants and recruiters are with light hearts. Convenient facilities multiply the amount of transactions and consequently enlarge the amount of rejections because the amount of job remains unchanged. From psychological and social perspectives, the infestation of rejections is not desirable.

On the other hand, even inconvenience requires users to make some efforts, its benefits are, as shown in section 2, psychologically desirable. Any examples shown in section 4 puts emphasize on subjective benefits that depend on individuals rather than on objective conveniences.

References

1. Kawakami, H.: Inconvenience utilizing Design. Kagaku-Dojin Publishing Company Inc. (2011) (in Japanese)
2. Nishigaki, T.: Fundamental informatics from life to society. NTT Pub. Co. Ltd. (2004) (in Japanese)
3. Kripke, S.A.: A Completeness Theorem in Modal Logic. The Journal of Symbolic Logic 24(1), 1–14 (1959)
4. Norman, D.A.: The design of everyday things. MIT Press (1998)
5. Norman, D.A.: Human-Centered Design Considered Harmful. Interactions 12(4), 14–19 (2005)
6. Norman, D.A.: Emotional Design. Basic Books (2005)
7. Norman, D.A.: Living with Complexity. MIT Press (2011)

8. Wilde, G.J.S.: Target Risk 2 –A new psychology of safety and health. PDE Publications (2001)
9. Taniguchi, T., et al.: Bibliobattle: an informal social interaction design mediated by book review. J. of Human Interface Society 12(4), 93–103 (2010) (in Japanese)
10. Suto, H., Okita, M.: Proposal of a Novel Communication Medium to Promote Community Using Homeless Cats. In: Proc. the 6th Workshop on Social intelligence Design, pp. 287–296 (2007)
11. Mizukoshi, S.: Media Biotope. Kinokuniya Publication Services (2005) (in Japanese)
12. Suto, H.: Media Biotope: Media designing analogous with Biotope. In: Proc. Int. Conf. on Computer Information Systems and Industrial Management Applications, pp. 75–80 (2010)

The Present Condition and Problems for Elderly People Participating in Communities

Noriyuki Kobayashi, Shohei Yamane, Kotaro Ohori, and Akihiko Obata

Fujitsu Laboratory Ltd.
{kobayashi.nryk,yamane.shohei,ohori.kotaro,
obata.akihiko}@jp.fujitsu.com

1 Introduction

The proportion of aged citizens in Japan has been growing and has reached the highest proportion in the world (23.1% in October, 2010 [1]). The time when we need to tackle the social welfare problem of an aging society has arrived. It has been reported that social welfare services are insufficient to reach all people who need help, resulting in 32000 lonely deaths in a single year [2]. It is difficult to solve the problem only by reforming formal care systems under conditions where the number of senior citizens is increasing and the number of younger citizens is decreasing. Community activity that supports elderly citizens is important when the proportion of aged citizens is increasing to prevent overreliance on formal care. The idea named PPK(pin-pin-korori [3]) thus emerged. This is the idea that the number of people requiring long-term care can be decreased by increasing the number of active senior citizens.

The following problems were reported in the annual report of the Cabinet Office of Japan for community activities for supporting elderly citizens [4].

- Difficulty for continuing community activities due decreasing numbers of participants in senior citizen clubs and increase of their age.
- Increase in elderly citizens who are isolated due to weak human connections with their regional community.

One of the approaches for solving these problems is promoting community activities that support senior citizens. ICT support is also expected to support such community activities. These mechanisms have been studied from the aspect of administration, however, there has been little research focusing on the view point of the senior citizen who actually participates in the community activities.

In this paper, we report an ethnographic study for finding out important design implications for ICT support for the community activities. The participants of the study had various problems and stories regarding senior citizens participating in community activities. We then discuss the implications of design for ICT support.

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 307–313, 2013.

2 Method

We conducted ethnographic interviews of fourteen senior citizens to find out about problem situations in their daily activities. Fourteen active seniors were selected as interviewees, since we could expect many tips from "lead users" [5].

In the ethnographic interviews, we asked interviewees to speak freely about their story of daily activities and community activities instead of answering predefined questions. Based on analysis of the transcriptions of the conversations, we collected various notes, such as episodes of participation, problems of participating in the community, surrounding situations, etc., from the story of the varieties of their activities. The collected notes were classified, and arranged into different types of problems and situations.

It is difficult to find out about episodes if you ask about the community activity directly. Therefore, we used the following interview sheets (Fig.1,2,3,4). Using the map sheets in Fig.1, we asked about the places where the interviewees visited daily. We then asked about the people who met at the places and their activities. Using the clock sheets in Fig.3, we asked about their daily activities from morning to night, and then clarified the daily activities that engaged in unconsciously. It is difficult to find out about daily activities because unconscious behavior does not leave an impression. However, it becomes possible to clarify these by looking back on the activities one at a time by using the interview sheets.

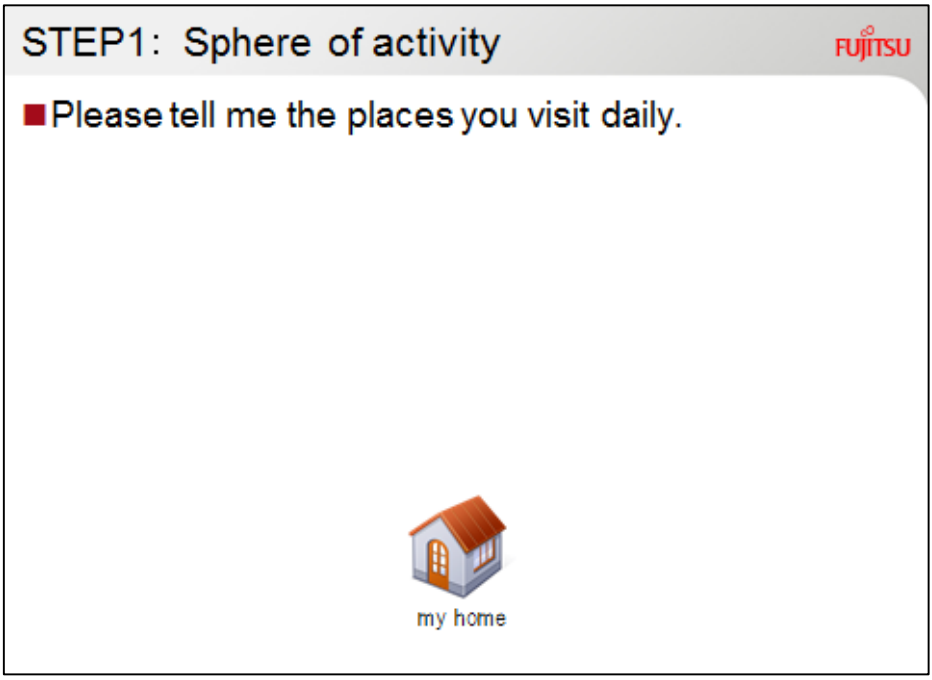


Fig. 1. Map sheet

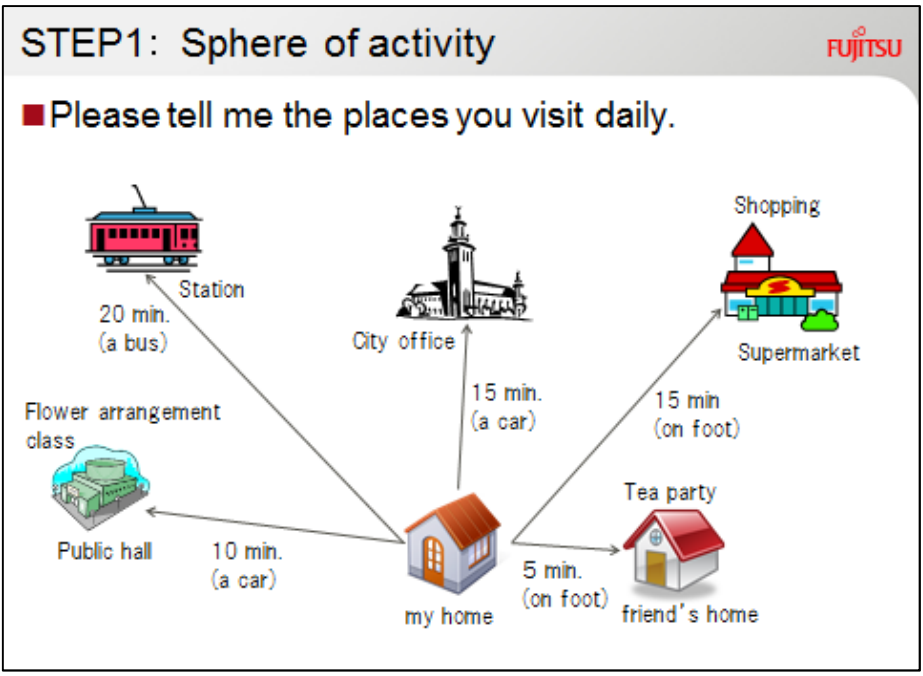


Fig. 2. Example of the entry in a map sheet

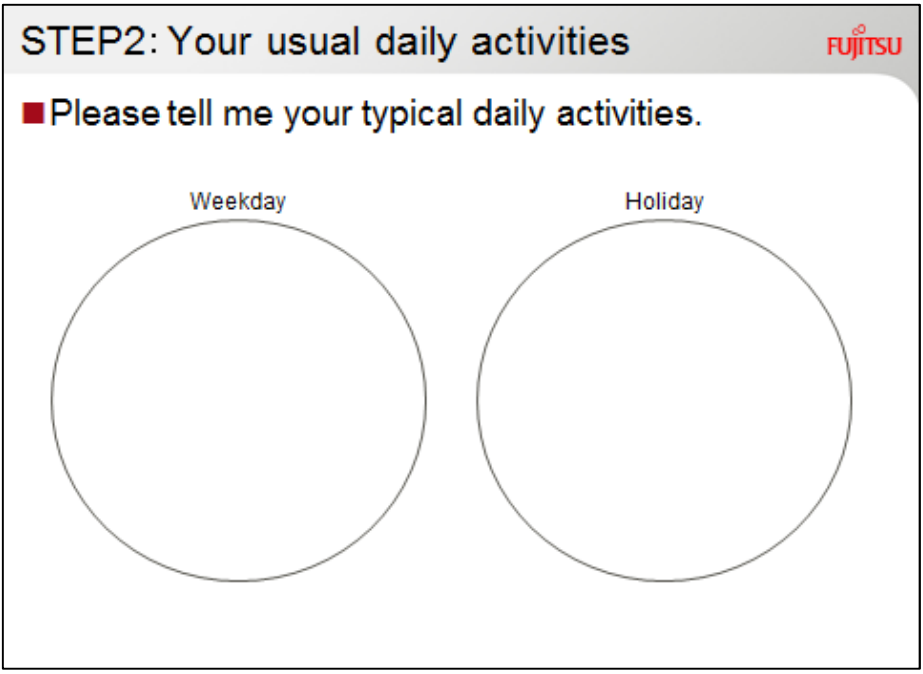


Fig. 3. Clock sheet

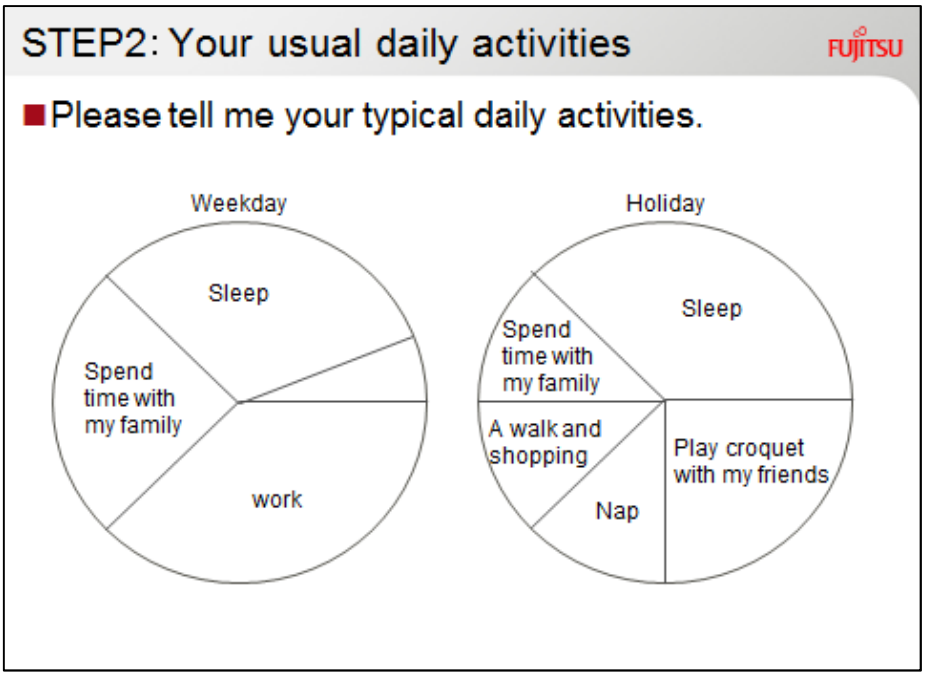


Fig. 4. Example of the entry in a clock sheet

3 Results

We got more than 100 episodes of problem situations for facilitating community activities from the fourteen interview transcriptions. The followings are example of episodes.

(1) A new community activity derives from an existing community activity.
We found that new communities were created triggered by community activities that already existed.

At the interview, we observed the following discussion.

"I joined the softball club because I was invited by a friend of the neighborhood association. The softball club is a very important community for me."

“I joined the sign language club because I have handicapped friends in another community”. These episodes might suggest that community activity itself acts as a trigger for a new community activity.

(2) Importance of participation by obligation
We found that social pressure was a one of the important factors for the participation of elderly people in communities.

"I was reluctant to join the neighborhood association. However, I have to take the role of the neighborhood council because it was my turn. But now, this community is very important for me"

"After I joined the neighborhood association, I created a baseball team with friends in the neighborhood association"

(3) Hesitation getting help from others

We found that a feeling of hesitation getting help from others is one of the reasons why elderly citizens refuse participation in community activities. An elderly person said "I feel uneasy about many people taking care of me free of charge in the community activities." "In a certain community, they reduce such hesitation by taking 100 yen as a registration fee."

(4) Barrier between generations

We found that there were few opportunities for meeting other generations in communities. Most regional communities are grouped by generation, such as boys and girls clubs, women's clubs, elderly citizens clubs, etc. There are few events for meeting different communities. An elderly person said "I have nothing to talk about with young men". "Our community has no events with other communities".

It is possible that a community in which most of the members are elderly will decline because younger people will not join. An elderly said, "There are no successors in our community. Who take cares of the community after we get older?" As a result, members who run the community were fixed, and there will be few chances for expanding their community.

On the other hand, we found a regional traditional festival event "O-maturi" in which different generations are participating fostered cross generational relationships. For example, we got episodes such as "I talked with a father of my friends", "I met a person who lives close to my house", etc.

4 Implications for ICT Design

As a result of the ethnographic interviews, we found that elderly people tended to not join unknown communities where there were no acquaintances. Therefore, it is necessary for the system to provide functions for encouraging them to join unknown communities. We designed a function that supports community members inviting elderly people. When an elderly person pushes the "NICE" button from the list of communities, members of the community get a message that tells them who showed interest in their community and the profile of that elderly person. The members of community can then decide who they should invite. We designed another function that supports elderly people joining a community with their friends. When an elderly person pushes the "NICE" button, the elderly person receives a message that shows their acquaintances registered as "Associate" who also showed interest in the same community.

As a result of interviews, after they joined a community out of obligation, they became interested in the community activities or they found other interesting communities. They might leave the community if they could not find interesting communities

and activities during the duty. Therefore, we designed a function that supports finding interesting communities. The system registers the elderly as an "Associate" when s/he joins the community. The system then supports finding a friend who has the same interest in the community by using the "NICE" button.

It was interesting for us that the interviewees felt guilty about getting kind help from others without giving back in return. Therefore, we designed a function that reduces the hesitation. The system supports selecting appropriate people for the task. When a community member inputs a task such as taking photos of the community, playing a guitar, etc., the system shows the appropriate persons based on the profiles of elderly people. It becomes easy to join a community by not requesting "please join us" but "please help us".

It is difficult to resolve the barrier among different generations by only focusing on supporting individual communities. We designed a new function supporting connections among different communities. A regional traditional festival could be a good trigger for fostering communication among different communities. The owner of the festival can call different communities to invite them to the festival using the new function. Members of each community can join a new community with their friends from their community. Invitation by community (not by individual) might reduce their hesitation of participating in a new community, and also might foster relationships among different generations of different communities.

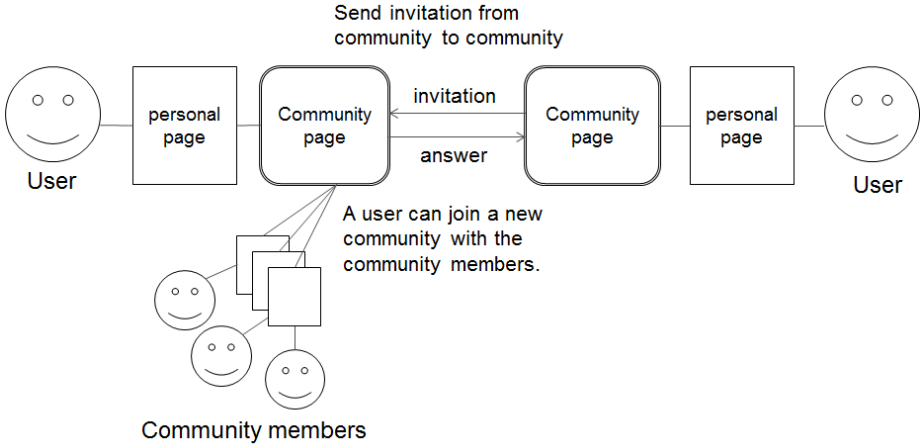


Fig. 5. Image of ICT design

The new function supporting connections with different communities might also help to activate and expand a community. Senior citizen communities tend to be run by fixed members and no new events are created. Owners of communities have to find people who help their events through their own network of acquaintances. The new function might help to find new activities and events of other communities, and easily get help for them using the new function.

5 Conclusion

We obtained design implications for community support based on interviews of community activities of active elderly citizens. However, these implications are still limited. We should investigate not only active elderly people but also non-active elderly people to get further implications. We will continue our ethnographic study on different types of citizens and other stakeholders and then conduct field experiments using an ICT prototype in some regions based on the findings of the field study.

References

1. Annual Report on the Aging Society: Cabinet Office Japan (2011), http://www8.cao.go.jp/kourei/whitepaper/index-w.html
2. NHK Muen Shakai Project, "Muen Shakai", Bungei SyunJyu (2010) (in Japanese)
3. Akiyama, H.: Science of aging society and social structure, Kagaku Iwanami Shoten (2010) (in Japanese)
4. A meeting for production of communities which elderly people etc. feel easy also alone and can live: Ministry of Health, Labor and Welfare Japan (2008), http://www.mhlw.go.jp/houdou/2008/03/h0328-8.html
5. Hippel, E.: Democratizing Innovation. MIT Press, Cambridge (2005)

Applying to Twitter Networks of a Community Extraction Method Using Intersection Graph and Semantic Analysis

Toshiya Kuramochi, Naoki Okada, Kyohei Tanikawa,
Yoshinori Hijikata, and Shogo Nishida

Graduate School of Engineering Science, Osaka University,
1-3 Machikaneyama, Toyonaka, Osaka, Japan
kuramochi@nishilab.sys.es.osaka-u.ac.jp,
{hijikata,nishida}@sys.es.osaka-u.ac.jp

Abstract. Many researchers have studied about complex networks such as the World Wide Web, social networks and the protein interaction network. One hot topic in this area is community detection. For example, in the WWW, the community shows a set of web pages about a certain topic. The community structure is unquestionably a key characteristic of complex networks. We have proposed the novel community extracting method. The method considers the overlaps between communities using the idea of the intersection graph. Additionally, we address the problem of edge inhomogeneity by weiting edges using content information. Finally, we conduct clustering based on modularity. In this paper, we evaluate our method through applying to real microblog networks.

Keywords: complex network, community extraction, intersection graph, hierarchical clustering, text mining, microblog network.

1 Introduction

Many researchers, having studied complex networks such as the World Wide Web, SNS networks, and the protein interaction network, have reported scale-free characteristics, the small-world effect, the property of high-clustering coefficient, and so on [1,2]. Recently, the community structure in complex networks is gaining increased attention from many researchers. The community structure shows the appearance of densely connected groups of nodes, with only sparse connections among groups. Many community detection methods have been proposed based on this definition [11]. They are applied to various complex networks. Communities in SNS networks shows a set of people with the same background or hobby. Communities in the WWW show sets of web pages related to a certain topic [4] and those in the protein interaction network show sets of proteins having the same function [6].

For community detection, researchers have started to show interest in whether overlaps between communities can be extracted [15,17,23]. The overlaps signify that one node belongs to several communities. For example, the Apple Inc. page

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 314–323, 2013.

is categorizable among computer category pages and audio category pages. It is important that a community detection method be able to assign a node not only to one community but also to several communities. Weights of all edges in complex networks are assumed to be the same in many community detection methods [19]. However, edges are rarely homogeneous in real networks. For example, various human connections such as those of businesses, hobbies, and organizations exist in SNS networks. It is important that the weights of these edges are not be treated as identical. They should be set individually. Many researchers conduct hierarchical clustering methods to complex networks for community detection. Most hierarchical clustering methods require advance input [3]: the number of clusters. However, the number of real communities is often unknown in real networks. Therefore, it is important that the number of proper clusters be decided automatically.

We have proposed a novel community detection method that can solve these three problems [8]. Our proposed method can extract overlaps between communities using the idea of the intersection graph. When there exist several subgraphs individual members are connected densely, we can make a new graph where the above subgraphs are converted to new nodes and edges are created when two subgraphs have common elements. Graph created by the above process is called intersection graph [10]. We also determine the weights of the edges in the intersection graph using two types of information: the degree of overlaps of the members and the similarity of content information such as text information and attribute information which appear in the network. Moreover, we use the hierarchical clustering method based on modularity proposed by Newman [11,13,14]. This method does not necessitate manual input of the number of clusters. In this paper, we evaluate general versatility of our method through applying to microblog networks.

2 Related Works

The problem of community detection in complex networks has been examined in various areas such as those of computer science and medical science [2].

Some researchers have attempted to extract communities in complex networks including the overlaps between communities. The overlaps mean that one node belongs to several communities. Everett et al. found them using the idea of the intersection graph [3]. Palla et al. found them by detecting cliques whose size was k and merging the cliques that shared $k-1$ nodes [15]. Fuzzy clustering, the method considers the notion of fuzziness and can assign one node to several communities is often used to extract overlaps of communities [17,23].

Our study weights edges in complex networks for dealing with edge inhomogeneity. Weighting edges in a network (usually a document network or hyperlink network) is popular in the area of the information retrieval. Some researchers improved the effectiveness of link analysis using content information. Jiang and Conrath measured the similarity between words using link information and semantic information of words [7]. Hung et al. improved the HITS algorithm by analyzing anchor text [5].

Many researchers use hierarchical clustering methods for detecting communities. The methods need input, which is the number of clusters preliminarily. Newman and Borgatti reported modularity as an indicator of how well the clusters are formed [12]. Newman proposed some clustering methods based on modularity [11,13,14]. These methods do not obviates manual input.

Our proposed method [8] considers the overlaps between communities using the idea of the intersection graph. Furthermore, we address the problem of edge's inhomogeneity by weighting edges using the degree of overlaps and the similarity of content information between sets (nodes of the intersection graph). Finally, we conduct a clustering method based on modularity, which does not necessitate manual input of the number of clusters. No study deal with all the above problems for detecting communities.

3 Proposed Method

In this section, we explain our proposed method [8]. The input of our proposed method is a graph of $G = (V, E)$, where V stands for the set of nodes and E signifies the set of edges. Additionally, content information is given to the nodes. We apply the following four steps to this graph.

Step 1. Enumeration of Dense Subgraphs: This method enumerates dense subgraphs (generally, they are called cliques) from an input graph of $G = (V, E)$.
Step 2. Conversion to the Intersection Graph: This method regards each subgraph enumerated in Step 1 as one new node and converts the input graph G to the intersection graph of $G' = (V', E')$.
Step 3. Calculation of the Weights of Edges: This method calculates the weights of edges E' in the intersection graph G' using the degree of overlaps and the similarity of content information between nodes V' (dense subgraphs) in G'.
Step 4. Clustering Based on Modularity: This method divides nodes V' into clusters using a clustering method based on modularity.

We applied the method of Everett et al. [3] to **Step 1** and **Step 2**. In **Step 1**, their method enumerates maximal cliques as dense subgraphs in an input graph of $G = (V, E)$. A clique is a subgraph in which an edge exists between any two nodes. Next, the method converts the input graph G into the intersection graph $G' = (V', E')$ in **Step 2**. Our method regards each dense subgraph enumerated in **Step 1** as one special node and makes the intersection graph $G' = (V', E')$ from the input graph $G = (V, E)$. When several sets (dense subgraphs) S_i $(i = 1, \cdots, n)$ are enumerated, our method generates a special node v'_i for each set S_i. If a common element exists in two arbitrary nodes v'_i and v'_j, then a special edge is put between them. The intersection graph is a new graph composed of special nodes and special edges [10]. When the method puts a special edge between special nodes, we can set the threshold of the number of common elements between the subgraphs corresponding to these special nodes. Finally, the method of Everett et al. conducts hierarchical clustering for the intersection graph. Our method address the edge inhomogeneity (in **Step 3**) and automatically detection of extracting communities number (in **Step 4**).

In **Step 3**, the proposed method calculate weights of edges in the intersection graph generated in **Step 2**. We use the degree of overlaps and the content information similarity between each subgraphs. There exist many types of measurement for presenting the degree of overlaps between X and Y ($d(X, Y)$) such as co-occurrence frequency, mutual information, Dice coefficient, Simpson coefficient and Jaccard coefficient [9,16]. For example, Jaccard coefficient is defined as:

$$d(X, Y) = \frac{|X \cap Y|}{|X \cup Y|}. \tag{1}$$

The proposed method uses vector space model [18] to calculate the similarity of the content information between two arbitrary sets X and Y. The method regards each set as one vector and calculates the *tf·idf* score for the keyword in the texts in the set. This *tf·idf* score becomes the element of the vector. Finally, the method calculates the similarity $sim(X, Y)$ between vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ corresponding to two sets X and Y using cosine similarity:

$$sim(X, Y) = \cos\theta = \frac{\boldsymbol{x} \cdot \boldsymbol{y}}{\| \boldsymbol{x} \| \| \boldsymbol{y} \|}. \tag{2}$$

Then, the proposed method calculates the weights $w(i, j)$ for the special edge between special nodes v'_i and v'_j (corresponding to set X and Y) using the degree of overlaps of sets $d(X, Y)$ and the similarity of content information $sim(X, Y)$. We can use several types of calculation function. In this work, we use function emphasizing the similarity of content information:

$$w(i, j) = w(X, Y) = \frac{d(X, Y)}{1 + \epsilon - sim(X, Y)}. \tag{3}$$

Here, ϵ $(0 < \epsilon < 1)$ is a constant used to keep the denominator from being 0.

Finally, in **Step 4**, the proposed method conducts clustering for community detection in the intersection graph. When a method extracts several clusters in a network, we must evaluate the currently detected clusters. Modularity is a broadly accepted indicator for evaluation. The indicator is simple and intuitive. Therefore, we adopt a clustering method based on the modularity that is proposed by Newman et al.[11,13,14]. When k clusters are given and P_k is defined as the sets of these clusters, the module function $Q(P_k)$ is the following.

$$Q(P_k) = \sum_i (e_{ii} - a_i^2) = \mathrm{Tr}(e) - |e^2| \tag{4}$$

$$\begin{cases} e_{ij} = \frac{1}{2m} \sum_{vw} A_{vw} \delta(c_v, i) \delta(c_w, j) \\ a_i = \frac{1}{2m} \sum_v k_v \delta(c_v, i) \end{cases}$$

4 Applying to Microblog Networks

We apply our method to real social networks. We select Twitter for this study which is the most popular microblog searvice. Twitter users post *tweets* (short

messages) and have conversations with other users through their tweets. If tweets have the word "@username," they are *mentions*—tweets for certain users. The word "*#hashtag*" in tweets means these tweets concern certain topics. Therefore, we can get the large amount of content information. In the Twitter network, all people can *follow* anyone without approvals, then the network contains links represent the unilateral interest. Links between users may represent relationships in the real world (university friend etc.) or that of interest (hobby friend etc.).

The purpose of this evaluation is to verify three questions:

- **Whether our method achieves better results than the conventional method:** We compare our method with the conventional method proposed by Everett et al. [3]. The conventional method converts an input graph into the intersection graph and conducts a simple hierarchical clustering.
- **Whether it is efficient to use content information for weighting edges:** Our method uses not only information about the degree of overlaps of sets but also the content information. We compare the method using both kinds of information with the method using only the degree of overlaps of sets. We confirm the effectiveness of the content information.
- **Whether the kinds of content information affect the results:** Twitter networks have several types of content information. We examine whether the results change according to the kind of content information.

4.1 Dataset

We make a dataset for the evaluation inviting test subjects who give true relationships between them and each member in the extracted communities. We followed users from a test subject to two in the radius (from the test subject up to the friends of the test subject's friends). Our experiment assume a situation extracting the communities of the real world or strong interest. There are some celebrities followed by million users in the Twitter network. A test subject may be not able to answer the proper relation between stranger users connected via such celebrities. We think it is important that all test subjects can answer the all relations between users in the dataset. Therefore, we set the threshold of the number of follow users and that of followers. Then, we removed such hub users (celebrities) in advance. Thresholds are set as 400 from prior study. Additionally, we collected profile texts and tweets (contain @usernames and #hashtags) as content information. The test subjects are 9 users who are all university students.

4.2 Implementation

Parameter Settings of the Proposed Method. We adopt the maximal clique as the dense subgraph in Step 1. We can use various sizes of the maximal clique (the clique threshold). If the clique threshold is 5, then the method uses only the maximal cliques that comprise more than four nodes. We set 3, 4 and 5 as the clique threshold. In Step 2, our method converts the original graph to the

Table 1. Statistical information of the dataset

method	content information
Prf	nouns in profile texts
Prf-H	nouns and #hashtags in profile texts
Prf-N	nouns and @usernames in profile texts
Prf-H+N	nouns, #hashtags and @usernames in profile texts
Twt	nouns in tweets
Twt-H	nouns and #hashtags in tweets
Twt-N	nouns and @usernames in tweets
Twt-H+N	nouns, #hashtags and @usernames in tweets

intersection graph. Here, it creates a special edge between two dense subgraphs (two special nodes) when they have common elements. We can set the threshold of the number of common elements (the overlap threshold) in this step. We set 2, 3 and 4 as the overlap threshold (In our prior work [8], we have found that the performance of the proposed method is very low when the overlap threshold is 1 under the influence of clustering method based on modularity). We conduct nine threshold conditions (clique threshold, overlap threshold) = $(3, 2)$, $(4, 2)$, $(4, 3)$, $(5, 2)$, $(5, 3)$ and $(5, 4)$.

In Step 3, we selected the Jaccard coefficient (eq. (1)) as the degree of overlaps of sets. We also selected profile texts and tweets as the content information. Our method extracts nouns, #hashtag and @username as keywords by conducting morphological analysis of the content information. The method calculates $tf \cdot idf$ scores for all keywords within one maximal clique (corresponding to a special node). The maximal clique can represent one vector. The similarity between maximal cliques is calculated using eq. (2). Finally, the weights between maximal cliques are calculated using eq. (3). We set $\epsilon = 0.1$ in this experiment. We use a greedy approach that repeatedly merges a pair of nodes to maximize the increment of the modularity [13].

Implementation of Community Extraction Method. As described in section 4.1, we use content information of two types: profile texts and tweets. We respectively designate the cases using profile texts and tweets as *Prf*, *Prf-H*, *Prf-N*, *Prf-H+N*, *Twt*, *Twt-H*, *Twt-N* and *Twt-H+N* (see Table 1). Hereinafter, we regard *Twt* as a representative example of these cases using the content information. We implemented the case using only the degree of the overlaps between sets (Jaccard coefficient) as the weights of edges in Step 3 to examine the contribution of content information analysis. We designate the case *NonCA* (without content analysis).

We implemented the method Everett et al. proposed [3] as a baseline method. The method comprises three steps. The first two steps of the method are the same as the first two steps (Step 1 and Step 2) of our method. Unlike our method, it uses a simple hierarchical clustering method in the third step. In detail, the method searches for a pair of special nodes that have maximal Jaccard

coefficient and merges the pair, repeatedly. We set the number of output clusters as the number of clusters including the test subject becomes the number of true communities (provided by each test subjects). We designate this case as *Everett's method*.

4.3 Evaluation Method of Extracted Clusters

Clusters of two kinds are extracted by a community extraction method: a cluster that includes the test subject and a cluster that does not contain the test subject. It is difficult for test subjects to judge the connection of members in the latter clusters. Therefore, we specifically addressed only those clusters containing the test subject. To measure the accuracy of the extracted clusters, we adopt the following evaluation process.

Step 1. A test subject enumerates all relation names for each person in the dataset (number of relation names is regarded as the number of true communities). The test subject can see the profile texts and latest 200 tweets of each person.
Step 2. We assume that the relation in the extracted cluster corresponds to each relation name. We calculate the precision, recall and$F$-measure per relation name for the cluster (For the calculation, we consider that the relation name is the name of correct relation for the cluster). The precision, recall and$F$-measure of a relation name N are calculated as follows.

$$\text{Precision}(N) = \frac{\#\text{ people whose relation name is } N \text{ in the extracted cluster}}{\#\text{ people in the extracted cluster}}$$

$$\text{Recall}(N) = \frac{\#\text{ people whose relation name is } N \text{ in the extracted cluster}}{\#\text{ people whose relation name is N in the dataset}}$$

$$F\text{-measure}(N) = \frac{2 \cdot \text{Precision}(N) \cdot \text{Recall}(N)}{\text{Precision}(N) + \text{Recall}(N)}$$

Step 3. We use the highest F-measure calculated in Step 2 among all relation names as the F-measure of the extracted cluster. We regard the relation name $\bar{N}$ that marked the highest F-measure as the relation of the cluster. We also use the precision and recall of a relation name $\bar{N}$ as the precision and recall of the clusters.
Step 4. We calculate the average values of the precision, recall and F-measure of all clusters. These values are regarded as the evaluation value of one test subject.

5 Evaluation of Extracted Clusters

We conducted an experiment of community extraction for clarifying three questions mentioned in Section 4. We show the average values of the precision, recall and F-measure in all 9 test subjects in the conventional method and our method in Figure 1-3.

We compare our method (*NonCA* and *Twt*) with the conventional method (*Everett's method*). In precision (Figure 1), the results of our method tend to become

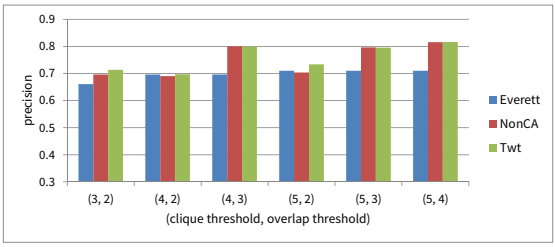


Fig. 1. Comparing the precision of Everett's method and our method

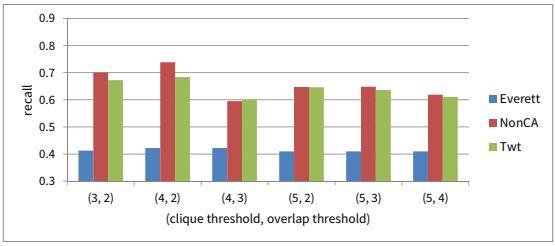


Fig. 2. Comparing the recall of Everett's method and our method

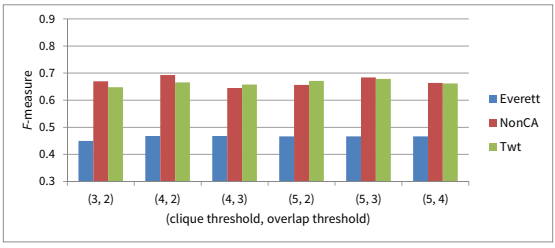


Fig. 3. Comparing the F-measure of Everett's method and our method

better when both thresholds become are larger. This is because users connecting strongly each other tend to survive while making the intersection graph. The proposed method overcomes the conventional method in the case when the overlap threshold become larger. In recall (Figure 2), the results of our method become better when both thresholds become smaller. Our method shows much higher recall than the conventional method in all conditions. At last, in F-measure (Figure 3), our method overcomes the conventional method in all conditions. Overall, we found our method brings a better result than the conventional method.

We compare *NonCA* and *Twt* to find out the effectiveness of content information analysis. In many threshold conditions, the precisions of *Twt* are greater than those of *NonCA* (Figure 1). Other hand, in recall, *NonCA* tends to overcome *Twt* (Figure 2). We cannot determine which method is better for extracting

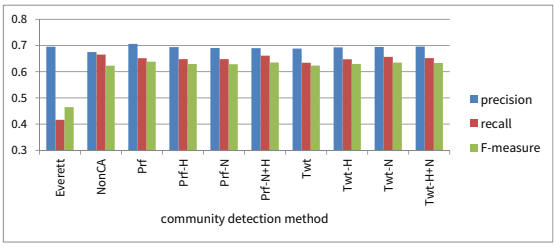


Fig. 4. Evaluation in condition $(4, 2)$ using content information of various kinds

communities because the difference in the F-measure is small (Figure 3). We think the reason is that text information in Twitter dataset contains many colloquial words and coined words. We conduct a simple morphological analysis that may be not enough effective to accurate extracting nouns from tweets.

Finally, we examine whether the results change according to the type of content information. We evalutate the cases using nouns, #hashtags, and @usernames in profile text and tweets as the content information in the threshold condition $(4, 2)$. We present results of *Everett's method*, *NonCA*, *Prf*, *Prf-H*, *Prf-N*, *Prf-H+N*, *Twt*, *Twt-H*, *Twt-N* and *Twt-H+N* in Figure 4. The results of cases using content information (all method without *Everett's method* and *NonCA*) are mutually similar. As we explained above, the influence to community detection of content information is not enough strong in our method (because of a simply morphological analysis). Additionally, many users in this experiment do not post tweets containing #hashtags or @usernames much. The volume of these information is not enough to influence the performance of our method.

6 Conclusion and Future Work

As described in this paper, we evaluated the proposed method for community detection [8] through applying to the Twitter dataset. We demonstrated the superiority of the proposed method compared to the conventional method. Although we compared the case using content information with the case not using that, we cannot show the advantage of proposed method. We think that is because the morphological analysis in our method is too simply to applying coined words and colloquial words in tweets and our experimental dataset do not contain much #hashtags and @usernames. As future workCwe will improve the morphological analysis and apply our method to other test subjects.

References

1. Albert, R., Barabasi, A.-L.: Statistical mechanics of complex networks. Rev. Mod. Phys. 74, 47–97 (2002)
2. Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D.U.: Complex Networks: Structure and Dynamics. Phys. Rep. 424(4–5), 175–308 (2006)

3. Everett, M.G., Borgatti, S.P.: Analyzing Clique Overlap. Connections 21(1), 49–61 (1998)
4. Flake, G.W., Lawrence, S., Giles, C.L., Coetzee, F.: Self-Organization of the Web and Identification of Communities. IEEE Computer 35(3), 66–71 (2002)
5. Hung, B.Q., Otsubo, M., Hijikata, Y., Nishida, S.: HITS Algorithm Improvement using Semantic Text Portion. WIAS 8(2), 149–164 (2010)
6. Huss, M., Holme, P.: Currency and commodity metabolites: Their identification and relation to the modularity of metabolic networks. IET Systems Biology 1(5), 280–285 (2006)
7. Jiang, J.J., Conrath, D.W.: Semantic Similarity Based on Corpus Statistics and Lexical Taxonomy. In: Proc. ROCLING 1997, pp. 19–33 (1997)
8. Kuramochi, T., Okada, N., Tanikawa, K., Hijikata, Y., Nishida, S.: Community Extracting Using Intersection Graph and Content Analysis in Complex Network. In: Proc. WI 2012, pp. 222–229 (2012)
9. Manning, C.D., Schütze, H.: Foundations of statistical natural language processing. MIT Press (2002)
10. McKee, T.A., McMorris, F.R.: Topics in Intersection Graph Theory, vol. 2. SIAM, Discrete Mathematics and Applications (1999)
11. Newman, M.E.J.: Detecting community structure in networks. Eur. Phys. J. B 38(2), 321–330 (2004)
12. Newman, M.E.J., Girvan, M.: Finding and evaluating community structure in networks. Phys. Rev. E 69(2) (2004)
13. Newman, M.E.J.: Fast algorithm for detecting community structure in networks. Phys. Rev. E 69(6) (2004)
14. Newman, M.E.J.: Finding community structure in networks using the eigenvectors of matrices. Phys. Rev. E 74 (2006)
15. Palla, G., Derenyi, I., Farkas, I., Vicsek, T.: Uncovering the overlapping community structure of complex networks in nature and society. Nature 435(7043), 814–818 (2005)
16. Rasmussen, E.: Clustering Algorithms, Information Retrieval: Data Structures and Algorithms. In: Frakes, W.B., Baeza-Yates, R. (eds.), pp. 419–442. Prentice-Hall (1992)
17. Reichardt, J., Bornholdt, S.: Statistical mechanics of community detection. Phys. Rev. E 74(1), 16110 (2006)
18. Salton, G., Wong, A., Yang, C.S.: A vector space model for automatic indexing. Communications ACM 18(11), 613–620 (1975)
19. Scott, J.: Social Network Analysis: A Handbook, 2nd edn. Sage Publications (2000)
20. Scripps, J., Tan, P.-N., Esfahanian, A.-H.: Node Roles and Community Structure in Networks. In: WebKDD/SNA-KDD 2007, pp.26–35 (2007)
21. Tasgin, M., Bingol, H.: Community Detection in Complex Networks using Genetic Algorithm. In: Proc. ECCS (2007)
22. Wasserman, S., Faust, K.: Social Network Analysis: Methods and Applications. Cambridge University Press (1994)
23. Zhang, S., Wang, R., Zhang, X.: Identification of overlapping community structure in complex networks using fuzzy c-means clustering. Physica A 374(1), 483–490 (2007)

Search Engine Accessibility for Low-Literate Users

Débora Maurmo Modesto, Simone Bacellar Leal Ferreira, and Aline Silva Alves

Departamento de Informática Aplicada, Universidade Federal do Estado do Rio de Janeiro, Rio de Janeiro, Brazil
{debora.modesto,simone,aline.alves}@uniriotec.br

Abstract. Search engines are often used to retrieve content on the Web, but it is not a simple activity for low-literate users since they have to know the technology and create strategies to query and navigate. Their interaction with search engines differ from high-literate users on strategies used, perception, communication and performance. In order to improve search engines and create solutions, we need to understand these users' needs. This research aimed to identify how search engine features influence the interaction of low-literate users. We analyzed the interaction of ten users through user tests that were part of a case study. Based on a limited set of features of a specific search engine, we identified what features were used, the perception about them and some barriers faced by these users. This study led to a list of recommendations for the development of search interfaces focused on low-literate users.

Keywords: Low-Literate Users, Search Engine Accessibility, Guidelines.

1 Introduction

The available tools on the Web are part of everyday life [1] and perform a social role especially for users with disabilities [2]. Therefore, there should be no barriers to access the Internet. There are users who have limitations related to literacy that can jeopardize the interaction mainly because the available content on the Web is mostly textual [3].

To help users to retrieve this content, search engines are often used: 80% of the access to Web pages comes from these tools [4]. Search involves analyzing different types of media, so it is a mentally exhausting activity that requires focus and attention [5]. Low-literate users have some limitations, particularly related to the strategies they use to conduct a search and the perception of interfaces [6-7]. Besides that, high-literate users interact with such tools in a different manner than low-literate users [3], [8].

In order to improve low-literate users experience on the Web, search engines should accomplish their mental models. Designers and developers should know how they interact with this kind of tools, what features are used and how they influence on user experience.

This research aimed to identify how search engine features influence the interaction of low-literate users. We observed how the interaction takes place and identified some barriers faced by these users. In order to do this, a case study was conducted

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 324–331, 2013.

leading us to a list of recommendations for the development of search interfaces focused on low-literate users.

Besides this introduction, this article is structured as follows: section 2 presents literacy concepts; section 3 presents related work; section 4 presents methodology; section 5 explains how the study was conducted; finally, section 6 presents conclusions.

2 Literacy in Brazil

According to United Nations Educational, Scientific and Cultural Organization (UNESCO), literacy can be analyzed by different perspectives and conceptualized in different ways. A common understanding about literacy involves oral, reading and writing skills, besides abilities with logic, mathematics, symbolic analysis (images and sounds) and text interpretation [9]. Nowadays it is a common approach to consider that these skills must be contextualized and they are not developed equally among different individuals. Besides, literacy concept also considers functional aspect that means the ability to apply oral, reading and writing on different areas of daily life, as in computing, ecology, health and other areas [9].

The concept of functional illiteracy varies from region to region. In Brazil, federal agencies as Brazilian Institute of Geography and Statistics (IBGE) adopt the same criteria as UNESCO to define function illiterates: people between 15 and 64 years old, which lack mastery of skills in reading, writing, calculations and science, corresponding to an education of less than four years of study [9-10]. In Brazil, it corresponds to incomplete 4th grade of elementary education. A research performed by IBGE showed that 29 million people in Brazil are functional illiterate that means more than 20% of the total population [10].

In this work, we adopt this concept but we call the participants as "*low literate users*", once one of the protocols used to guide the study recommended not to call them functional illiterates [8].

There are other criteria that could be used on the study instead of considering only age and years of study on formal education. There are institutions that developed specific literacy and numeracy tests, with levels of difficulty and punctuation, to evaluate people's skills in reading, writing, calculations and science. However, these tests are private, so there are no public tools that allow us to classify users this way [8].

3 Related Work

3.1 Taxonomy of Web Search

Web search can be categorized in three groups: (a) informational, when users want to find more information about a topic; (b) transactional, when users want to perform operations after searching, as downloading a file; (c) navigational, when users want to find a site from some organization they already know [4], [11-12].

3.2 General Users' Behavior

Some studies address general users' behavior when using search engines: On informational or transactional queries, users usually focus on the title and description of a search result. Images are also an expected type of media, but videos can distract users once it is not possible to comprehend its full meaning quickly [4]. On navigational searches, users tend to ignore results from the fourth position on and they focus also on the URL [4].

If users do not find what they want on first results, they tend to perform another search [13]. Users are usually influenced by suggested results' relevance, identified through positioning [14-15].

Another study identified that experience influence users' behavior. Novice users adopt search strategies less flexible than expert ones, have difficulty to formulate a query and do not know how search engines work [16]. Another study stated that only a few users know how to use advanced features [17].

A study about terms suggestion showed that most users preferred to refine the query manually [18]. On other study some people stated that this feature was helpful but it was also a distraction. However, after a week of log analysis, users started to use it in an iterative way [19].

Users usually do not know how the search engine works. It creates incorrect expectations about outcomes. Explanations presented about how search engine works helps users to understand outcomes [20].

Some studies showed that grouping results by categories is a good way to present them [21-22]. Present a good description of the outcomes is also helpful. Some studies showed that summarization of the page and highlighting terms were good to improve result analysis [23-24].

3.3 Low-Literate Users' Behavior

Low literate users usually do not check correctness of information [6]. These users usually cannot scan results as high literate users do [4], so they need to read every word to understand information [6]. They tend to have more difficulty recovering from errors or changing search strategy. They also become confused when navigating in pages full of information and links. Their performance is usually worst when compared to high literate users, since they take longer to finish tasks, spend more time on the same page and visit more pages [6].

A study stated that they need features that support decision about results such as in which link they should click or if information is relevant or not. The same study mentions that this kind of feature is more important than the ones to support them to define query [25].

No study was found about how search features influence low-literate users. This paper shows how search features affect low-literate users' experience, mainly Brazilians whose first language is Portuguese.

4 Methodology

This was a qualitative and exploratory research. The research took place in Rio de Janeiro, Brazil, and consisted of a case study to analyze the interaction of ten low-literate users with Google search engine through user tests. The participants were between 15 and 64 years old and had less than four years of study on formal education [9]. They were all Brazilians from different regions of the country.

This approach can be considered limited because it does not consider extracurricular capabilities and years of study are related to a feasible educational goal by Brazilian government [4]. If other criteria were used to select participants, maybe other results could be found.

Tests were operationalized through "Protocol for conducting usability testing with a focus on accessibility" [26], that defined steps to accomplish planning, preparing, conducting and reporting results, and "Protocols for Web accessibility evaluation involving functional illiterates" [8], that details the approach and execution of the user tests.

People were recruited in schools with youth and adults education classes, churches and residential buildings. As recommended by the protocol [8], a portable usability lab was set and used on the tests that were performed in various locations to facilitate transportation for participants.

A questionnaire was applied in order to find more information about users' profile. Based on this, Google was chosen for observation once it was the main search engine used for all participants. Besides that, search volume on this search engine corresponds about to 66% of all searches on the Web [27-29]. In 2011, Google Brazil (www.google.com.br) reached more than 92% of searches performed by users in Brazil [30].

The search engine was explored on its default state and all features were available with no customization. A limited set of features related to activities such as writing, reading and formulating search, guidance, navigation and feedback were selected for analysis. This set consisted of: auto complete, spell checker, related searches, advanced search, filters, layout (header, search bar, advertisements, pagination and footer), search results, keyboard navigation, "I'm Feeling Lucky" button, and Google Instant features (as page and result preview). Some aspects such as simplicity of the text, amount of terms used, how they elaborate a query, perception and orientation were also analyzed.

The case study consisted of two units of analysis. On the first unit five users interacted with Google search engine to perform five tasks that varied by difficulty level. The first three tasks were considered easy and it was expected that all participants concluded all of them to finish the test. The fourth task was considered a medium task since it involved notions of magnitude, as maximum and minimum and advanced vocabulary. The fifth one was considered a difficult task since it involved notions of history, current events and interpretation skills. Tasks also varied by search goals, since three of them were informational tasks, one was navigational and one transactional. Test could finish after five tasks concluded or thirty minutes, what have happened earlier. After the test, each user was interviewed and answered questions about perception of the features´ utility.

On the second unit, five other users interacted with two resources that were not used on the first unit of analysis by anyone, but were considered useful: filters and advanced search. Before performing each task of this unit, a video was shown teaching how to use each feature.

Textual material used in the study, including the questionnaire, tasks and informed consent, were prepared with the assistance of a checklist for plain writing for Web [31]. Data was analyzed following Four-Phase Framework for Search [32]. This framework states that every search consists of four main phases that are formulation, action, review of results and refinement. Features were grouped on these categories and analyzed according to each phase´s goals.

5 Case Study

Three men and two women participated of each unit of analysis. Data collected on questionnaires indicated that eight users were less than two years of experience with internet and two were less than five years. Despite the wide range of age adopted as criteria, we selected users who had similar education and experience with computers in order to minimize a possible bias caused by age difference. Five participants said they usually ask for help when using search engines. Eight participants indicated that the main difficulty is to know whether a word is spelled correctly. The result analysis was also considered a difficult task by six participants. On the other hand, seven participants stated they were comfortable when they need to formulate a query. Regarding to search topics, products were the most mentioned (cited by eight participants), followed by music and videos (both mentioned by seven participants).

5.1 Recommendations for Search Engine Development Focusing on Low-Literate Users

Below are presented the recommendations developed based on observation of low-literate participants during user tests:

1. Provide features that help writing and problem formulation (like "spellchecker", "autocomplete" and "related searches"). Show changes made by the "spellchecker" as soon as the results are presented, positioned above the results, so they are quickly visualized. Present results corrected by "spellchecker" and provide feedback, indicating that terms were changed. Show terms suggestion to complete the query while user types next to the search field to enable quick viewing. Show the feature "related searches" below the results for easier query refinement.
2. Provide "filters" not only for refinement, but also for formulating the query. In both cases, filters should be formatted as categories or as a menu. Place "filters" in areas with less emphasis, such as the header of the page. Present results related to "filters" or categories in a different format from conventional results.
3. Provide a large text box to write the query. The terms that the user typed should not be hidden so he does not forget the words he used and do not get confused about the research problem.

4. Low-literate users cannot handle too much information at the same time and get confused with lots of text, so display around seven main results at a time (at least five and at most nine). In general, people feel more comfortable to handle this amount of results [33]. Secondary outcomes that lead to internal pages of a website can be displayed once they not hinder the understanding of users, but should not be excessive.
5. Present the results divided into pages and display a paging feature to navigate between them, positioned at the end of the results page. Also use the expression "see more" beyond the page numbers, because this term is more familiar to users.
6. Provide a visual indication of results' relevance, which is not only showed by the page rank and positioning on the page. Positioning has not a clear meaning for these users.
7. Do not present other media formats or filtered results among the conventional results because it confuses users. Allocate a page area to present this kind of outcome.
8. Display the title and description for a result emphasizing the first one. Show other information as the URL on demand, only if user requests. Generally, low-literate users do not visualize this information and do not use it to decide whether to click on a result. Search terms should be highlighted on the description, in order to keep user focused on the subject of task. Distinguish snippets extracted from different parts of a website through background colors, for example. Avoid using suspension points for that.
9. Features that provide instant feedbacks are recommended, however, should be prominently displayed so they can be readily seen.
10. Provide features to help users to decide whether or not to select a result. These inputs should be showed only on demand. For example, "page preview" feature provides inputs to the user to decide whether to select a result, but it's not used by low-literate users once they cannot comprehend what is in the page without reading carefully its content.
11. Advanced features such as "keyboard navigation", can be available if the interface is also used by advanced users. However, low-literate users do not make use of these resources.
12. The footer area of the results page is less visualized. Provide information that does not need emphasis in this area.
13. Use tips and directions about the use of the interface so that it does not distract users nor overload the page with lots of information.

6 Conclusion

During this research, a set of recommendations was developed, considering low-literate users' preferences, such as writing instead of reading and features that don´t impact performance. They also addressed issues related to user's needs such as: (1) features to assist them in writing and formulating search problem, (2) minimize distractions, (3) provide clear and visible feedback, (4) provide instructions about search

engine behavior, (5) facilitate recovery from errors, (6) provide ways to stay focused on one activity at a time, (7) decrease the amount of text and results, (8) organize search results by categories.

These recommendations still need a validation since it was not addressed on this study. Despite the similarity with other search engines, generalization should be carried carefully once no tests on other tools were performed to validate the guidelines at this time. These recommendations can assist developers in creating interfaces for search engines or search features within websites. It is expected that interaction of low-literate users is enhanced on this kind of tools and they find information more easily.

References

1. Nielsen, J., Loranger, H.: Usabilidade na Web: Projetando Websites com Qualidade. Elsevier, Rio de Janeiro (2007)
2. Ferreira, S.B.L., Nunes, R.R.: e-Usabilidade. LTC, Rio de Janeiro (2008)
3. Kodagoda, N., Wong, B.: Effects of Low & High Literacy on User Performance in Information Search and Retrieval. In: Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction, pp. 173–181. The British Computer Society, Swinton (2008)
4. Thurow, S., Musica, N.: When Search Meets Web Usability. New Riders, Berkeley (2009)
5. Hearst, M.A.: Search User Interfaces. Cambridge University Press, New York (2009)
6. Kodagoda, N., Kahan, N., Wong, W.: Identifying Information Seeking Behaviors of Low and High Literacy Users: Combined Cognitive Task Analysis. In: Proceedings of NDM9, the 9th International Conference on Naturalistic Decision Making, pp. 347–354. The British Computer Society, London (2009)
7. Gupta, N.K., Rosé, C.P.: Understanding Instructional Support Needs of Emerging Internet Users for Web-based Information Seeking. Journal of Educational Data Mining 2, 38–82 (2010)
8. Capra, E.P.: Protocolos para Avaliação da Acessibilidade Web com a Participação de Analfabetos Funcionais. Master Dissertation. Universidade Federal do Estado do Rio de Janeiro, Rio de Janeiro (2011)
9. United Nations Educational, Scientific and Cultural Organization.: Understandings of Literacy. In: The Education For All Global Monitoring Report: Literacy for Life, pp. 149–159 (2006), http://www.unesco.org/education/GMR2006/full/chapt6_eng.pdf
10. Instituto Brasileiro de Geografia e Estatística: Síntese de Indicadores Sociais: Uma Análise das Condições de Vida da População Brasileira. (2010), http://www.ibge.gov.br/home/estatistica/populacao/condicaodevida/indicadoresminimos/sinteseindicsociais2010/SIS_2010.pdf
11. Broder, A.A.: Taxonomy of Web Search. ACM SIGIR Forum 36(2), 3–10 (2002)
12. Manning, C., Raghavan, P., Schütze, H.: Introduction to Information Retrieval. Cambridge University Press, Stanford (2008)
13. Spink, A., Jansen, B.J.: A Study of Web Search Trends. Webology 1, (2004), http://webology.ir/2004/v1n2/a4.html
14. Keane, M.T., O'Brien, M., Smyth, B.: Are People Biased in Their Use of Search Engines? Communications of the ACM 51, 49–52 (2008)
15. Bar-Ilan, J., Keenoy, K., Levene, M., et al.: Presentation Bias is Significant in Determining User Preference for Search Results: A User Study. Journal of the American Society for Information Science and Technology 60, 135–149 (2009)

16. Hölscher, C., Gerhard, S.: Web Search Behavior of Internet Experts and Newbies. Computer Networks 33(1-6), 337–346 (2000)
17. Machill, M., Neuberger, C., Schweiger, W., Wirth, W.: Navigating the Internet: A Study of German Language Search Engines. European Journal of Communication 19, 321–347 (2004)
18. Anick, P.: Using Terminological Feedback for Web Search Refinement: A Log-based Study. In: Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Informaion Retrieval (SIGIR 2003), Toronto, Canada, pp. 88–95 (2003)
19. Anick, P., Kantamneni, R.G.: A Longitudinal Study of Real-time Search Assistance Adoption. In: Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2008), Singapore, pp. 701–702 (2008)
20. Muramatsu, J., Pratt, W.: Transparent Queries: Investigating Users' Mental Models of Search Engines. In: Proceedings of the Twenty-fourth International ACM Conference on Research and Development in Information Retrieval, New Orleans, LA (2001)
21. Käki, M.: Enhancing Web Search Result Access with Automatic Categorization. Ph.D. dissertation, University of Tampere, Finland (2005)
22. Käki, M.: Findex: Search Result Categories Help Users When Document Ranking Fails. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2005), Portland, Oregon, USA, pp. 131–140 (2005)
23. Tombros, A., Sanderson, M.: Advantages of Query Biased Summaries in Information Retrieval. In: Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 1998), New York, pp. 2–10 (1998)
24. White, R.W., Jose, J., Ruthven, I.: A Task-oriented Study on the Influencing Effects of Query-biased Summarization in Web Searching. Information Processing and Management 39(5), 707–733 (2003)
25. Gupta, N.K., Rosé, C.P.: A Foray into Understanding the Next Billion Search Users (2010b), http://www.cs.cmu.edu/~nkgupta/papers/chi2010a_submitted.pdf
26. Henry, S.: Just Ask: Integrating Accessibility Throughout Design (2007), http://www.uiaccess.com/accessucd/
27. Experian. Search Engine Analysis, http://www.hitwise.com/us/datacenter/main/dashboard-23984.html
28. Comscore. comScore Releases U.S. Search Engine Rankings (January 2012), http://www.comscore.com/Press_Events/Press_Releases/2012/2/comScore_Releases_January_2012_U.S._Search_Engine_Rankings
29. Netmarketshare. Desktop Search Engine Market Share, http://netmarketshare.com/search-engine-market-share.aspx?qprid=4&qpcustomd=0
30. Serasa Experian. Share Mensal de Buscas do Google Atinge 92,15% em Janeiro, http://www.serasaexperian.com.br/release/noticias/2011/noticia_00381.htm
31. Barboza, E., Nunes, E.: A Inteligibilidade dos Websites Governamentais Brasileiros e o Acesso para Usuários com Baixo Nível de Escolaridade. Inclusão Social 2(2), 19–33 (2007)
32. Shneiderman, B., Byrd, D., Croft, W.B.: Clarifying Search: a User-Interface Framework for Text Searches. D-Lib Magazine, 1–15 (1997), http://www.dlib.org/dlib/january97/retrieval/01shneiderman.html#formulation
33. Miller, G.A.: The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information. Psychological Review 101(2), 343–352 (1955)

Identifying and Representing Elements of Local Contexts in Namibia

Kasper Rodil[1], Kasper Løvborg Jensen[2],
Matthias Rehm[1], and Heike Winschiers-Theophilus[2]

[1] Dep. of Architecture, Design and Media Technology
Aalborg University
Denmark
{kr,matthias}@create.aau.dk
[2] Mobile Future Lab and Dep. Software Eng.
Polytechnic of Namibia
Namibia
leafcastle@acm.org, heikew@polytechnic.edu.na

Abstract. In an attempt to represent local context in a 3D visualisationfor rural elders in Namibia we have foundmajor differences in theconceptualizationof this context between external and local partners in the co-creation process. Through the evaluation of a mobile context capture tool we found a clear disconnection of community members with both abstract and absolute representations of points, paths and areas. From this we discuss how the local concepts of space and time as frames of reference can not be represented adequately with our current selection of contextual data, and how we are engaging in participatory activities to derive a common understanding of contextual representations.

Keywords: context, indigenous knowledge, Participatory Design, context-aware, re-contextualization.

1 Introduction

Rural Herero communities in Eastern Namibia have for many years been self-contained in terms of transfer of local knowledge across generations. Children and youth have been listening to their elders, participating in chores and practical work in the villages, thereby sustaining an often tacit and uncodifiable knowledge. In Wenger's terms it could be considered a 'repertoire' of indigenous knowledge in 'communities of practice' [1]. These rural communities perform actions within their own cultural context. There are obvious local benefits in transferring knowledge on husbandry, herbal lore etc. and the tacit knowledge transferred through intra-personal interaction effectively adds to preserving local culture, customs and traditions. Due to formal education the youths are increasingly detached from their cultural traditions and context. The majority of the youths attend schools often far away from the villages, where many only return on holidays. The curriculum is compliant with international standards inconsiderate of local conditions. Valuable indigenous knowledge is lost day by day,

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 332–341, 2013.

and future generations could very well suffer from this absence of cultural roots and traditions, such as lack of self-awareness, cultural adaptation and self-worth. Moreover a major risk is the impact on the ecosystem in which the Herero tribe has lived sustainably since their settlement in Namibia in the 17th century.

Since 2008 we have been in close collaboration with a group of Elders (knowledgeable men respected in their community) in a village in Eastern Namibia. The overall objective of the project is to preserve local knowledge, but also find ways to transfer parts of that knowledge to de-situated youths from the region. Early in the process we have come to terms with the different world views of us, primarily Western trained designers, and local Herero elders. We have emphasized a dialogical approach with intensive collaboration and co-design. Our methodological stand is within Participatory Action Research (PAR).

In the ethos of PAR lies the acknowledgment by designers that they are from the onset limited in their understanding of users and the users' context. As reported by Nielsen et al.(2003) [2] one of the principles of PAR is mutual learning. Co-designers (in this case Elders) acquire technological skills meanwhile designers attempts to fine tune their sensitivity to the new setting and the skills and knowledge within that domain. The differences are in particular revealed when the participants (co-designers and designers alike) have their origin in different epistemologies. As Participatory Designers we attempt to understand the context through a shared perspective. Limitations will unavoidably occur, but we must investigate were these cut-offs manifest in the design of context-aware systems. As we have reported earlier, the foundations on which we design interfaces from might have little use in cross-cultural collaborations [3], [4]. Thus one of our approaches has been to substitute traditional (Western) interface metaphors with localized metaphors. Yet we hypothesize that it may still be desirable to explore these differences in the interaction design to overcome HCI issues of both textual and computer literacy and unnatural interface conventions such as folders, menus and files born out of the earlier computing days.

2 What Is a Context?

The meaning and use of context in computing systems and HCI has been debated for decades. Abowd et al. (1999) provided an operational definition of context within the scope of context-aware computing as: "Context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves."[5].

When introducing context-awareness, Schilit et al. (1994) alluded that the notion of context is dynamic and goes beyond static concepts such as location [6], e.g. lighting, noise level, network connectivity, communication costs, communication bandwidth and even the social situation is part of the context. Thus a context is also defined by people and their actions, at least if we include people interacting with artifacts. Dourish (2004) describes context as arising from activity:

"Context isn't just 'there', but is actively produced, maintained and enacted in the course of the activity at hand" [7]. Even just from these three perspectives on 'context' we can establish that the characteristics and description of any context is highly complex. We can consider it to be a temporal state of a changing space influenced by those that occupy and interact with and within it.

Thus it is impossible to capture and represent an entire context or expect to experience the exact same state once again. It is optimistic to think that a system is able to capture the complexity of an 'interwoven' situation of actions performed in space and time, as we always make a choice of which aspects of a context we capture. As Grudin (2001) explains on capturing context digitally:

"The context that is captured is removed from its context, namely the context that is not captured." [8].

This implies varying degrees of implications for artifacts using 'some' of the information available from the context. An application that address users based on identity, their activities and their location might not require extra information from the context they are in to achieve the set objectives. For instance, a mobile application that identifies a user then pulls the user's characteristics, activity and location to then push a suggestion of the route to a preferred dining place. As described by Baldauf et al. (2007) there are a variety of sensors available to further interpret the context (e.g. biosensors, thermometers, cameras etc.)[9]. Though it is of high practical value in mapping and modeling efforts, but the arguably "objective" snapshots of the perceivable space-time continuum often offers a poor reflection of the messy, subjective world of cultural, historical and spiritual meanings attributed to objects, places and situations by human beings. The objective of capturing a context in order to represent it digitally while actually producing a new context is a delicate matter. Especially when engaged in cross-cultural design of artifacts and preservation of cultural knowledge. Instead of taking the pessimistic stance of not being able to capture the whole context, we could examine the affecting elements defining and capturing a context.

3 Context Discussed through a Case Study

Since 2010 we have investigated the potential of 3D visualizations as means for preserving local knowledge. The main driver for the 3D visualization approach is to capture the local physical context into a visible medium unrestricted from requirements of local literacy and being transformable into digital experiences as virtual worlds or serious games. A cornerstone in the project is to enable rural elders to capture landmarks and objects of significant importance (houses, fenced areas, water holes etc.) at the research site through tagging with external sensors. This data set will represent the contextual backbone of recreating the village as a 3D representation for de-situated youths to explore.

An example of use could be that the youths take an online virtual tour of the village while listening to their grandfathers telling locally collected stories

at places represented in a virtual context yet separated in time and space from the recording. A cardinal point is whether the realism of sensor-captured GPS coordinates and compass bearings can translate into contextually "accurate" experiences. Thus it is fundamental to the research of representing a context, that we can pin-point some of the challenges to facilitate knowledge preservation and knowledge transfer. In this article we refrain from discussing the limitations of 3D visualization. As will be highlighted later, the action of representing a context through an incomplete perspective into a different form is not without implications.

4 CARACAL: A Tool to Capture Context

CARACAL is a context-aware tool designed for mapping and tagging objects and places in the field using sensor-enabled mobile devices [4]. Initially the tool was developed as an aid to the researchers of the project when engaged with re-mapping the spatial layout of the village. It was iteratively designed and implemented over a series of field trips and tuned to optimize the efficiency for the researcher.

Fig.1 shows the flow of interaction from the prototype. As can be seen the main screen allows quick access to capturing the core types of data points and gives feedback to the user as data is entered. CARACAL logs spatial information as either points(e.g. objects), paths(e.g. walking routes) or polygons (e.g. homesteads), which can each be tagged with additional information. It also allows for the capture of rich media through pictures, video or audio which is automatically Geo-tagged using the GPS, electronic compass and further allows entry of meta data by the user.

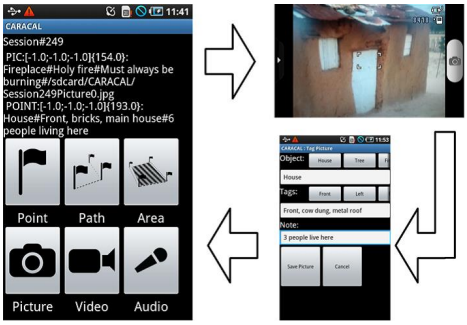


Fig. 1. The figure shows the CARACAL interface and flow of interaction

While a study showed that this approach was feasible for researchers and that we could indeed gain a lot in terms of efficiency and better data sets, we also found that there were several important shortcomings. Firstly, it would never

scale to the level we would want to use it due to the shear amount of information to be captured and tagged in even a relatively small village. Secondly, we need to capture the "right" data, i.e. the information that has meaning and is important to the local users. It was thus evident that local participants needed to be included and that the optimal solution would be a tool that would enable *them* to capture and map out their own environment according to their local perspective.

We conducted an exploratory field study of the application with a number of participants from different groups in the community and got mixed results.

As it turned out, the best performers were the youngest participants as they were more receptive to the instructions and fast to learn to interact with the application. The elders experienced much more challenges as they had a harder time with the touch screen based phone.

The biggest problems were however of more conceptual nature. It was evident that there were plenty of opportunities to enter 'erroneous' data because the original application was developed from the researchers' understanding and mental model of what the data should be used for. For example, this resulted in wrong Geo-tagging and or missing information about the captured pictures.

We also conducted an unsupervised field trial with a single user traveling to his home village in the North of Namibia. He used the application to track walks in the area while listening to an elder from his local community. The user is an IT master student and fairly proficient with mobile technology, so the fact that he did not report any serious problems with the application was not a surprise. The most interesting result from this part was the data itself, showing a spatio-temporal flow through the environment and context as they walked around. The data was not comprehensive enough that we could actually recreate the environment but simply overlaying it onto a satellite map in Google Earth gave an insight into that space (see Fig.2). What we also found was that the story behind the walk and the points tagged on the way were missing. Thus a limitation of the data is the lack of narration and re-contextualization.

Fig. 2. The figure shows the student's and elder's walk

5 Discussion of Capturing and Representing the Physical Domain

5.1 Spatial

The coordinates obtained from the CARACAL are supposed to be used for instantiating objects at the 'right' places and with a Cartesian mapping in a 3D world. But as reported in the evaluation of the interface:"The key findings were however, that the conceptual idea of 'points' and underlying purpose of 'tagging' was not well understood by the participants." [4]. The participants (elders) were not: "distinguishing between taking a picture of a landmark from a distance and tagging the very same from close proximity" [4].
Thus within the evaluation of the interface we began questioning if coordinates perceived to be objective were indeed objective, or if local concepts of space are different to the Western concept of mapping a space. Pursuing our conceptualization and representation carries the risk to override a tacit cultural way of understanding space. This might not have severe implications in many other application domains, but within our project the very objective is to represent cultural and local knowledge. Research suggests that there indeed might be a difference in cognition of space across cultures. Haun et al. (2011) describe studies investigating cultural differences of spatial cognition and spatial language between Dutch children and children from the Akhoe Hai//om hunter-gatherer group of Northern Namibia [10]. The Dutch children have a dominant ego-centric cognitive frame of reference, where the participating Akhoe Hai//om children have demonstrated a dominant Geo-centric (absolute) frame of reference. The former preference is to imagine the coordinate system centered in the ego. The latter is where the coordinate system is absolute with cardinal directions (N, E, S, W). This has an impact on how people orient themselves in the world, thus if a mapping into a new 3D geometric space is constructed, the mapping might need re-alignment according to local concepts.
Thus the action of capturing absolute 'points' on a global grid presents the scaled distance between objects and activities as a fixed and measurable length. A fundamental question is whether there are local differences in perception of distance between points. While we have not conducted formal studies as of yet, we ask ourselves if activities, paths or places within the village are locally perceived to be of a different scale? This could relate to differences in considered significance of paths over places, or vice versa. While it remains a speculation for now, we believe that in order to capture a space we must make sure we also represent the distances. Based on some conversations with community members we believe that a local perception of distance is that objects within a homestead are considered being almost at the same place, whereas everything situated outside is considered far away. Thus we attempt to acknowledge this in our representation.

5.2 Temporal

Another, yet unresolved, parameter to be represented is the concept of time. While traditional context-aware systems are easily able to log time of events,

time between events etc. the fundamental question is: do Western academics have a different perception of time than the co-designers and the end-user group? Should time be used as a convenient separation of activities (as in meeting calendars, time tables etc.) or is activity and time even connected and to what extent? Research shows that local concepts of time can have different characteristics. Janca and Bullen (2003) explain that: "the Aboriginal concept of time differs from the Judeo- Christian perception of time in that Aboriginal people do not perceive time as an exclusively 'linear' concept (i.e. past–present–future) and often place events in a 'circular' pattern of time according to which an individual is in the centre of 'time-circles'...". [11]. We have experienced that time is perceived differently in the village than e.g. any random day at our academic institutions. Sometimes we wait hours for the participants in the village although having agreed to meet at a certain place at a specific time. As we put ourselves under pressure to collect data during the relatively short village stays, we often struggle accepting delays relative to our own plans. However, we must remember that our research interferes with their daily routines and activities. It is also important to note that while we can consider ending a meeting due to another meeting pressing on in the calendar, they might not end theirs because there are still important matters to talk about. While time can be a fix point for activities we have experienced that social activities in the village are often defining time in the village. From a methodologically perspective we try to adapt to the local 'rhythm' of doing things. It shows us how we perceive time as a separation of activity. Practically it means that when we design databases, we should be cautious to taxonomically appoint time as the separating factor of events/activities. If we assume a re-contextualization to be truthful we should be careful not to override local perspectives with Western conceptualizations.

The purpose of the examples here is not only to elucidate a potential difference in concepts, but to stress the divergence of the underlying framework of reference. Thus if we as designers presuppose something to be a fact or common ground we might lose sensitivity to the fact that it might not be universal.

6 Participatory Design for Synthesizing the Context

When discussing the use of context in computer systems Dey (2001) provides a description of how humans have some success in transferring concepts to each other:

"Humans are quite successful at conveying ideas to each other and reacting appropriately. This is due to many factors: the richness of the language they share, the common understanding of how the world works, and an implicit understanding of everyday situations." [12].

Most of us are familiar with the saying: "it must be understood in the context that...". This statement implies that individuals in dialogue indeed are talking about the same state if sharing similar perspectives. Yet people having shared experiences can argue about what has happened, or the meaning of something although they were both part of the same event. This seemingly trivial example

on subjectivity is much more prevalent in a cross-cultural project due to the obvious background difference between parties. Differences in language, world-views, value beliefs and a collaboration containing two distinct cultural traits as either collective or individual oriented leads to substantial divergence of perception and desirable representations.

In a classical example Quine (1960, Pp. 29) analyses the difficulties of establishing a common ground and a common understanding about the context of the conversation, when you do not speak the same language [13]. His example is called the Gavagai problem: Imagine walking in the bush with a Herero guide and suddenly he points at a rabbit and says "Gavagai". Quinn's argument about the indeterminacy of the situation (and possible translation of the utterance) takes into account that it is possible to derive several conflicting representations based on the observable data, e.g. Gavagai could mean rabbit, or dinner, or actually the bush under which the rabbit is sitting. Additionally, he states that there is no way to know which of the competing interpretations is the correct one, which poses a slightly ontological problem, when you assume that an objective representation (some kind of ground truth) exists (see [14] for an in-depth discussion).

From these examples it might be too ambitious to expect parties to have a similar notion of the context under dissection. Thus, there will consistently be a mismatch on interpretation and representation. While seemingly impossible to resolve we have tried to understand the differing perspectives through Mutema (2003, Pp.5), who explains that the goal for interpretation is by fusing horizons:

"The inter-subjective nature of the research process allows for the researcher's interpretations to be checked, reinterpreted and evaluated by the actors." [15].

This perspective on understanding difference is emphasized in our participatory design approach. The arguably objective nature of the shared interpretation is in the hands of the parties involved. In this example we could explain it like: we attempt to increase our sensitivity to interpret an unfamiliar context through dialogue. Meanwhile the participants acquire skills on technology and design in order to be critical towards the shared representation. We have reached a stage in the project where the elders are indeed critical towards some of the prototypes. But we have experienced in many cases that at the point of critique new knowledge about the context is brought forward. The following example is from recently published work and highlight the value of the local dialogue [3]. We considered the concept of the trash can desktop icon as not being universally understandable. We re-constructed this icon into looking like a hole in the ground which we saw a local researcher from the village used to dispose garbage in. In the evaluation of this metaphor we were made aware by the elders that they do not throw away, but keep those parts for later use. Thus without the local evaluation of an interface action we would probably not have learned about that particular view point of not throwing away. Needless to say which consequence that misinterpretation would have for a knowledge transfer tool preserving cultural viewpoints. The misrepresentation is *our* context unaligned conceptualization. They suggested us to implement that the objects targeted for deletion should

be moved back into the menu from where it came. We argue that the next design iteration being informed from this evaluation is not a representation from either part, but a product of fusing horizons. Although we remain cautious to any impact new concepts might have on their context.

7 Conclusion

We have argued, given the application domain of transferring indigenous knowledge and local culture, that designers should be wary that transfer of partial information from one context to another might have implications on the development and research objectives. We have argued through experiences with a Herero community in Namibia, that perceived value-free and objective measurements might produce a distortion and that those measurements should be investigated further. That interpretation as an activity is in the hands of all actors –participants and designers.Our proposed solution is participation and inclusion in the design process through dialogue.

Acknowledgements. We wish to thank the community members for their commitment and friendships.

References

1. Wenger, E.: Communities of practice: Learning, meaning, and identity. Cambridge University Press, Cambridge (1998)
2. Nielsen, J., Dirckinck-Holmfeld, L., Danielsen, O.: Dialogue Design-With Mutual Learning as Guiding Principle. International Journal of Human-Computer Interaction 15(1), 21–40 (2003)
3. Rodil, K., Theophilus, H.W., Jensen, K.L., Rehm, M.: Homestead creator: a tool for indigenous designers. In: Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design, NordiCHI 2012, pp. 627–630. ACM, New York (2012)
4. Jensen, K.L., Theophilus, H.W., Rodil, K.: Tapping into local lore: toward scalable local mapping and tagging for rural Africa using mobile devices. In: Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design, NordiCHI 2012, pp. 631–634. ACM, New York (2012)
5. Abowd, G.D., Dey, A.K.: Towards a better understanding of context and context-awareness. In: Gellersen, H.-W. (ed.) HUC 1999. LNCS, vol. 1707, pp. 304–307. Springer, Heidelberg (1999)
6. Schilit, B., Adams, N., Want, R.: Context-aware computing applications. In: First Workshop on Mobile Computing Systems and Applications, WMCSA 1994, pp. 85–90. IEEE (1994)
7. Dourish, P.: What we talk about when we talk about context. Personal and Ubiquitous Computing 8(1) (2004)
8. Grudin, J.: Desituating action: Digital representation of context. Human–Computer Interaction 16(2-4) (2001)
9. Baldauf, M., Dustdar, S., Rosenberg, F.: A survey on context-aware systems. International Journal of Ad Hoc and Ubiquitous Computing 2(4) (2007)

10. Haun, D., Rapold, C.J., Janzen, G., Levinson, S.C.: Plasticity of human spatial cognition: Spatial language and cognition covary across cultures. Cognition (2011)
11. Janca, A., Bullen, C.: The Aboriginal concept of time and its mental health implications. Australasian Psychiatry 11(s1), S40–S44 (2003)
12. Dey, A.K.: Understanding and using context. Personal and Ubiquitous Computing 5(1), 4–7 (2001)
13. Quine, W.V.O.: Word and Object. MIT Press, Cambridge (1960)
14. Raatikainen, P.: On how to avoid the indeterminacy of translation. The Southern Journal of Philosophy 43(3), 395–413 (2005)
15. Mutema, G.: Phenomenology, hermeneutics and the study of indigenous knowledge systems. Indilinga: African Journal of Indigenous Knowledge Systems 2(1) (2003)

A Framework for Community-Oriented Mobile Interaction Design in Emerging Regions

Monica Sebillo[1], Genoveffa Tortora[1], Giuliana Vitiello[1], Pasquale Di Giovanni[1,2], and Marco Romano[3]

[1] Department of Management and Information Technology, University of Salerno, Italy
{msebillo,tortora,gvitiello}@unisa.it
[2] School of Computer Science and Informatics, University College Dublin, Ireland
pasquale.di-giovanni@ucdconnect.ie
[3]Universidad Carlos III de Madrid, Spain
mromano@inf.uc3m.es

Abstract. The interaction design of mobile applications for a specific community of users, requires a deep knowledge of that community in order to establish appropriate usability requirements and tune the subsequent development activities. This is especially true when the target community is situated in one of the developing countries. There, mobile devices are increasingly playing the role that personal computers play in 'developed' countries, in spite of quite different cultural, social and technological constraints. In this paper we propose a 'community-centered' design approach, where different aspects of a community are used to formulate usability goals and are taken into account throughout the design, prototyping and testing phases, so as to ensure the actual application deployment by users of that community. The case study of a project carried out with a community of farmers in Sri Lanka is used to illustrate the approach.

Keywords: Analysis and design methods, Human Centered Design and User Centered Design, Human Factors Engineering Approach, Interaction design

1 Introduction

In the last few years several design guidelines and hci patterns have been issued to provide guidance in the design of mobile applications addressing common usability challenges [10-12, 15-18]. For instance, knowledge of cultural norms of reading conventions and how people process information is used to suggest how to design elements for an interface and place items on it so that users will understand. All such guidelines and patterns share the basic assumption that mobile users are everywhere, want to access information quickly and to be able to manipulate it easily. Mobile users are in the design focus and the adoption of common best design practices is recommended to guarantee usability with users who keep moving around with their devices and whose focus on the screen can be frequently distracted by the surrounding environment.

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 342–351, 2013.

However, the design of mobile applications for a specific community of users, requires a deep knowledge of that community, that often goes beyond the single user and introduces cultural and social constraints which may invalidate some of the general guidelines. Knowledge of the community, in that case, helps establishing appropriate usability requirements and tuning the subsequent development activities on the community itself. This is especially true when the target community is situated in one of the developing countries [5-8]. There, mobile devices are increasingly playing the role that personal computers play in the developed countries in spite of personal, social and technological constraints.

In this paper we introduce a 'community-centered' design approach, where the social, cultural, technological and economic aspects of a community are used to formulate adequate usability goals and are taken into account throughout the design, prototyping and testing phases, so as to ensure the actual deployment of the application by users of that community (Fig. 1). The paper is organized as follows. In Section 2 we describe the two-layered ethnographic observation method which characterizes the proposed approach. We provide guidelines for community profiling in the mobile world that can be used to separate the 'once-for-all' community observation activity from the specific problem domain analysis. In Section 3 we illustrate the community-oriented design approach on a concrete case study, i.e., the design of a mobile application for a community of farmers in Sri Lanka.

2 Guidelines for Community Profiling in the Mobile World

The human/sociological/behavioral aspects which characterize a community of people are usually derived from the observation of their Cultural, Social and Technological contexts [9]. In the following, a set of guidelines is defined to be used for profiling a community of people using mobile devices, independently of the specific problem domain.

- *Social Context* (Social Organization, Ethical beliefs etc.)
  - Consider the social organization of the community. It "describes the collection of values, norms, processes and behavior patterns within a community that organize, facilitate and constrain the interactions among community members"[1]
    - Find out the social necessities and limitations.
    - Find out social relationships between individual subjects belonging to the same community.
    - Find out possible participation in governative or voluntary organizations.
- *Cultural Context* (Language, Education)
  - Consider the average cultural level of the target community. The designed interface should be easily understood by any stakeholder.
  - Consider the average education level of community members. Never make a design choice based on incorrect knowledge assumptions about the user.

- Different languages may affect the use of text in the visual design of the interface. This is especially true for communities living in some Eastern countries and where the official spoken language uses an alphabet different from the common Latin alphabet. Consider that:
  - sentences could be hard to represent on small screens,
  - the (virtual) keyboard could be missing some characters of the language alphabet. This would again affect the choice of the mobile device.
- In some countries more than one language is spoken. So, consider the necessity to design a multi-language interface.
- Consider the semiotics of the target community. Signs, colors, symbols, metaphors can have different meanings in different environments.

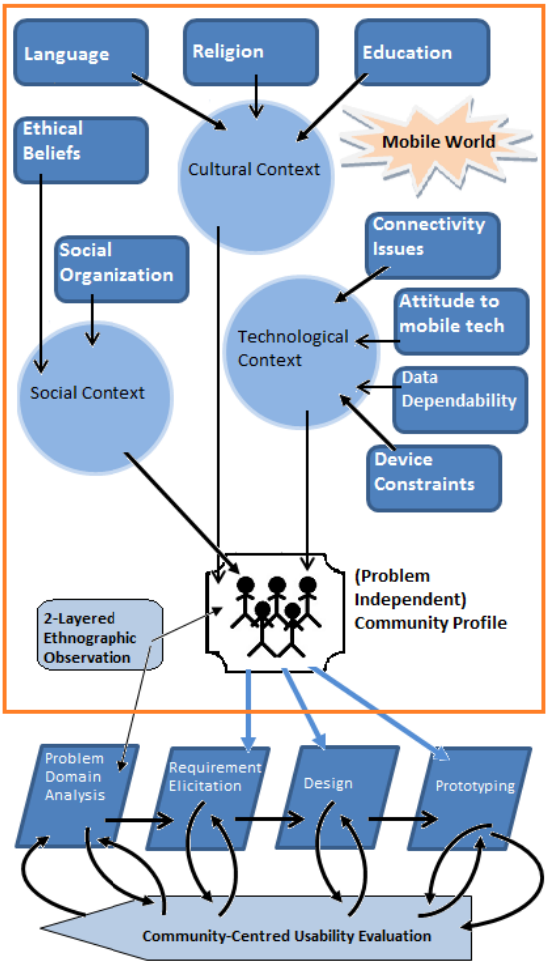


Fig. 1. Community-oriented design and development

- *Technological Context* (Available technology, Familiarity with mobile devices)
  - Consider the technological means available in the geographic area of the community. Also consider the average degree of familiarity with the mobile technology you are planning to use and the attitude to learning new technologies.
  - Mobile devices operations often depend on remote services. Therefore connectivity issues are paramount in this context:
    - consider the data quantity that the system needs to transmit on the wireless networks
    - bad connections can cause loss of reliability and can make the application progressively slower causing usability problems too.
  - The application should run on the majority of the devices available in the community. The world market trends can suggest some devices, however in specific communities particular technological ecosystems could be found.
  - Consider the device models available to that community in the specific country. Advanced devices may be present in rich communities (but this is not a rule), in some countries some models could not be available.

2.1 Problem Domain Analysis

Differently from the traditional approach, the analysis of current practices will aim at observing people from the target community within scenarios related to the specific problem domain [14]. Therefore, being aware of the common characteristics of the community of users and of their general needs with respect to the mobile world will allow designers to focus on the following points:

1. *what* is expected from the target mobile application
   - what tasks are currently performed?
   - what roles?
   - what artifacts?
   - what interactions?

2. *why* users need the application
   - motivations for adopting a mobile application supporting current activities

3. what the *temporal context* of interaction is
   - when is each task accomplished?
   - when should the system be deployed?

4. what the *geographical context* of interaction is
   - where do the observed activities take place?

Requirements elicitation will come out of the answers to the above questions. In the following section we describe the community-oriented design methodology applied on a concrete case study.

3 The Case Study of SLN4MOP

The Social Life Networks for the Middle of the Pyramid (SLN4MoP) is an International Collaborative research program started in 2011 that aims to provide real-time information to support activities related to livelihood delivered by mobile phone applications targeted to meet the needs of people in developing countries [19].The goal of SLN4MOP project is to leverage the enormous reach of mobile phones equipped with myriads of sensors such as GPS and camera, to develop the next generation of social networks that not only connect people to people, but also have the capability of providing real-time, context-sensitive local information by aggregating data from a variety of sources. In order to move towards that goal, Sri Lanka was chosen as the country where a pilot research study could start.

In this section we show how our approach can be adopted for the community of Sri Lankan people who are willing to use mobile phones to improve their daily working activities. In particular, the case study deals with the design of a mobile application meant to support farmers in the initial phases of crop selection and planning.

3.1 Profiling the Community of Sri Lankan People

The Sri Lankan society presents opposing aspects: people reflect both some of the typical facets of the modern advanced western societies and some aspects strongly connected to a multiethnic culture, full of ancient traditions. As a result of the initial ethnographic study, we were able to depict a general community profile for people living in Sri Lanka, who may benefit from the use of mobile devices for their daily activities [2, 3]. This represents the common knowledge about the social, cultural and technological contexts, which could be exploited for several application domains.

In the following we present some relevant aspects of the community profile, which directly impacted some of the design choices made for the farming application.

Cultural Context - Sri Lanka is a real multicultural nation. The community is made of two main ethnic groups, namely the Sinhalese and the Tamil. 83% of the members speak Sinhalese and the remaining 17% speak Tamil. However, English is the third official language, mainly spoken in the cities.

*Social Contex*t - Young people have not been extremely influenced by modernity and the effects of westernization. 80% of them have religious beliefs, the majority being Buddhists. Mobile devices are quite widespread but there are still barriers to their adoption, especially among parents who are often concerned with security and reputation issues. Moreover, people's general attitude to ICT is often influenced by the opinion of prestigious members of the community, such as local temple priests.

Technological context - Like in many developing societies, a significant gap in ICT (digital divide) can be observed in Sri Lanka, especially among people living in villages, sometimes even missing electric power supply. In order to encourage the adoption of technology and support the diffusion of culture, the local government has recently created the so called "Nenasala telecenters" supply [4]. Computer desktops and laptops are absent for the majority of individuals, while mobile devices are widespread enough. Statistics say that 86.5% of the population owns at least one SIM

(Subscriber Identity Module). Nevertheless, even if the Sri Lankan mobile network currently covers almost the totality of the urban areas, temporary lacks of connectivity, may affect services which heavily rely on network uptime.

In this context, the use of ICT and especially mobiles, have the potential to reduce the information and knowledge gaps in the affected areas.

3.2 Problem Domain Analysis: The Cultivation in Sri Lanka

Agriculture in this developing economy employs the largest share of the workforce; yet it contributes the least to GDP (Gross Domestic Product) when compared to the Industry and Services sectors. The cultivation pattern of Sri Lanka is bi-modal. This is mainly due to the two monsoons, which bring rainfall to the island in two distinctively different periods resulting in two distinct cultivation seasons [13].

From the survey we conducted with local farmers, it was noticed that due to this distinction there are specific crops to be cultivated in each season. Other factors that contribute to the crop selection are type of soil, weather, water, financial status, pest and diseases. Although several factors may affect the selection of crops for cultivation, farmers are used to choose crops primarily based on the profit. As a result they are biased by the market price of the previous year and they tend to select crops that produce a high yield within a short time period.

None of the interviewed farmers had a specific land extend per crop in mind at the time of cultivation, though they have a better understanding about the different types of crops that can be grown on a particular season. This knowledge is gained mainly based on the practical experience that came via ancestors. Moreover, during the study it was further identified that farmer awareness with respect to what others grow is limited to the neighboring farmers. This is mainly gained by watching what others are growing. They were reluctant to get this information by formal communication due to social phobia and the competition among farmers.

Different selling mechanisms are adopted by farmers in Sri Lanka. Some bring the harvest directly to the market, while around 90% depend on a middle person, namely the transport agent or the shop keeper. However, none of them get help from the government to sell the harvest. Another interesting fact is the behavior with respect to the selling prices. The selling price is a dynamic value which changes very frequently at a particular market. The farmers reported that they were often unable to predict the price as it changed vigorously within few hours. However, it was revealed that, according to their experience, they were unable to gain a good price for their harvest at the market, when all farmers tend to grow the same crop at the same time. As a result, in some cases farmers gave up farming and joined other industries to find their living. If this trend would continue in the future, it might have a huge impact for the economic growth of the country.

In the following, we summarize the most important claims about the domain specific issues observed from the field study:

- users are disposed to use some technological instruments provided that they are not invasive;

- governmental centers aimed at supporting agricultural activities are located all around the farms;
- there is a very low level of trust among the members of the same community, insomuch as not sharing basic information about their crop production.

3.3 The Community-Oriented Design

The user interface and application functionalities have been designed on the basis of the social, cultural and technical contexts characterizing the Community Profile of Sri Lankan people. Based on that, we understood that the adoption of smartphones was paramount in this work, since users are distrustful of technology with the exception of mobile devices, which are widespread. Moreover, farmers may need training that could be provided by experts of governmental centers.

The application was also expected to provide support to a critical task on the basis of data updated directly by users. Therefore, it was paramount to augment the perception of the data reliability.

The social and cultural contexts characterizing the profile of Sri Lankan community, besides the existing technological constraints, led us to choose the Android operating system for our mobile application, due to the high variety of device models running that system that are sold in that country. The technical and cultural issues, along with the small display size of the devices on which it will be executed also deeply influenced the choices regarding the design of the application user interface.

Major technical limitations are represented by the low resolution of the screen and the aforementioned limited sizes, the low computing power and the limited amount of main memory available. Such limitations, as well as the necessity to preserve as much as possible the battery life, were a determining factor in our decision to limit the amount of data processed directly on the smartphone. However, the need to communicate, via the mobile network, with a remote server has not been a trivial challenge and a lot of efforts have been made to handle with the payload of our application on the network. Even if the Sri Lankan mobile network currently covers almost the totality of the urban areas, we had the necessity to find the right tradeoff with regard to the data exchange modality and the quantity of data transmitted, while limiting as much as possible the impact that a temporary lack of connectivity, typical of a GSM / UMTS network could have on the general usability of our system.

More specifically, in order to decrease the total payload of exchanged network data we adopted two main solutions. First of all, we chose JSON (JavaScript Object Notation) as a lightweight data-interchange format. JSON essentially allowed us to present the data choosing the desired level of detail, with less verbosity than XML.

The second solution we adopted is based on a geographical information filter. Data sent from the remote server are preventively filtered on the basis of the farmers' location. The server collects information about products adequate for the soil type and climatic conditions of the region of interest and about the products which have been selected by neighbors so far. Such information is then aggregated, combined and computed in order to send only relevant data through the network. It is worth to point out that the given solutions addressed a general implementation problem characterizing

the target community and may therefore be adopted in different domains without further efforts. The same consideration applies to the visual interaction paradigm we decided to adopt, based on colors and images with little textual descriptions.

From the social context we knew that even if English is spoken among the members of our target community, users feel more comfortable using their own language so we decided to design our interface by allowing multiple languages. At the very beginning we designed the interface for the English language and then we translated the text in Sinhalese. We noticed that the graphical representation of this language is often significantly longer than the English text representation. This caused a visualization problem and a resulting redesign of some screens of the application that made the interface auto-adapting to the text dimensions.

At a later stage, the knowledge of the social phobia among farmers, who are reluctant to share with others the information on what crops they are growing, led us to design the interface so that the needed information could be provided anonymously and visualized using a 'traffic light' metaphor. Figure 2 shows the crops catalog filtered by the farmer's geographic position. For a better readability, a zoomed portion of the screen is also displayed. The colored background of each crop indicates the approximate quantity already in production. Figure 3 shows the comparison screen. A farmer can use this feature if he wishes to get more information about a specific subset of crops.

Fig. 2. The Crops Catalog

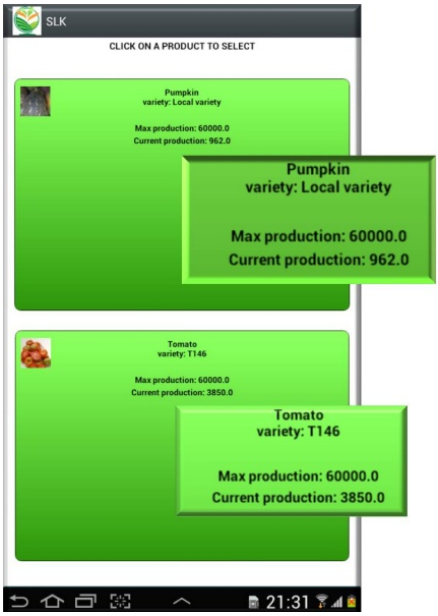


Fig. 3. Crops Comparison

4 Conclusion

The development of mobile applications is paramount to support users living in developing countries to improve their lives. However, a deep understanding of the target community is crucial for the realization of an application that is effectively accepted by those users. In this paper we propose a community-centered mobile design methodology, based on the idea that relevant aspects of the community can constitute a bulk of knowledge upon which mobile interaction designers can build appropriate solutions to a specific problem. The case study of Sri Lankan farmers presented in the paper is part of a wider international project, which is meant to provide ICT support to people from developing countries, with the goal to improve their quality of life. In the near future we plan to validate the proposed methodology designing other mobile applications for users who belong to the same socio/cultural community. This could be the case, for instance, of an application meant to support fishermen of Sri Lanka in their trading activities.

References

1. Mancini, J.A., Martin, J.A., Bowen, G.L.: Community capacity. In: Gullotta, T.P., Bloom, M. (eds.) Encyclopedia of Primary Prevention and Health Promotion, pp. 319–330. Kluwer Academic, Plenum, New York (2003)

2. Di Giovanni, P., Romano, M., Sebillo, M., Tortora, G., Vitiello, G., Ginige, T., De Silva, L., Goonethilaka, J., Wikramanayake, G., Ginige, A.: User Centered Scenario Based Approach for Developing Mobile Interfaces for Social Life Networks. In: Proceedings of the 2012 First International Workshop on Usability and Accessibility Focused Requirements Engineering (UsARE 2012), Zurich, Switzerland, June 4, pp. 18–24 (2012) ISBN 978-1-4673-1846-4 © 2012 IEEE
3. Di Giovanni, P., Romano, M., Sebillo, M., Tortora, G., Vitiello, G., De Silva, L., Goonethilaka, J., Wikramanayake, G., Ginige, T., Ginige, A.: Building Social Life Networks through Mobile Interfaces–the Case Study of Sri Lanka Farmers. In: Proceedings of the IX Conference of the Italian Chapter of AIS, Rome, September 28-29 (2012)
4. Telecenters in Sri Lanka: The Nenasala project. Technical Report. Skill International Private Limited, Lakshman Building, Colombo
5. Medhi, I., Patnaik, S., Brunskill, E., Gautama, S.N.N., Thies, W., Toyama, K.: Designing mobile interfaces for novice and low-literacy users. ACM Trans. Comput. Hum. Interaction 18(1), Article 2 (2011)
6. Danis, C., Ellis, J., Kellog, W., Hoefman, B., van Beijima, H., Daniels, S., Loggers, J.: Mobile phones for health education in the developing world: SMS as a user interface. ACM DEV, London (2010)
7. Putnam, C., Rose, E., Walton, R., Kolko, B.: Mobile phone users in Kyrgyzstan: a case study of identifying user requirements for diverse users. In: Professional Communication Conference (IPCC). IEEE (2009)
8. Jayaweera, P., Senaratne, R.: Mobile service portal for rural fisher community development. In: Proceedings of the International Conference on Advances in ICT for Emerging Regions (ICTer), pp. 66–70. IEEE, Colombo (2011)
9. Dey, A.K.: Understanding and Using Context. Personal and Ubiquitous Computing 5(1), 4–7 (2001)
10. Preece, J., Rogers, Y., Sharp, H.: Interaction Design, 3rd edn. John Wiley and Sons (2011)
11. Mayhew, D.J.: The usability engineering lifecycle. Morgan Kaufmann Publishers (1999)
12. Longoria, R.: Designing Software for the Mobile Context, 1st edn., vol. XVII, 152 p. 72. Springer, Longoria (2004)
13. Central Bank of Sri Lanka Annual Report (2010)
14. Rosson, M.B., Carrol, J.M.: Scenario-based design, The human-computer interaction handbook: fundamentals, evolving technologies and emerging applications. Lawrence Erlbaum Associates Inc., Mahwah (2002)
15. Hoober, S., Berkman, E.: Designing Mobile Interfaces. O'Reilly Media (2011)
16. Neil, T.: Mobile Design Pattern Gallery. O'Reilly Media (2011)
17. Di Chiara, G., Paolino, L., Romano, M., Sebillo, M., Tortora, G., Vitiello, G., Ginige, A.: The Framy User Interface for Visually-Impaired Users. In: Proc.s of Int. Conf. on Digital Information Management, ICDIM 2011, pp. 36–41. IEEE Press, Melbourne (2011)
18. Ginige, A., Romano, M., Sebillo, M., Vitiello, G., Di Giovanni, P.: Spatial data and mobile applications - general solutions for interface design. In: Tortora, G., Levialdi, S., Tucci, M. (eds.) Proceedings of the International Working Conference on Advanced Visual Interfaces, Capri Island, Naples, Maggio 21-25, pp. 189–196 (2012)
19. Social Life Networks for the Middle of the Pyramid, International project, `http://www.sln4mop.org/sln/home/index/home`

The Effects of (Social) Media on Revolutions – Perspectives from Egypt and the Arab Spring

Christian Sturm[1] and Hossam Amer[2]

[1] Hamm-Lippstadt University of Applied Sciences
christian.sturm@hshl.de
[2] The German University in Cairo, Egypt
hossam.amer12@gmail.com

Abstract. The Arab Spring has been titled as "social media revolution". While there is no doubt that Facebook and twitter contributed significantly to the course of events, its role can only be understood when put into a broader technological and historical context. Therefore, we looked at the role of technology during prior revolutions such as the French Revolution in 1789. It turns out that media technology has played an important supportive role in social and political movements throughout history. The changing characteristic, however, has always been the speed of information diffusion. Therefore, it is concluded that social media should be seen as supportive but not exclusively responsible for the events of the Arab Spring.

Keywords: Social media, society, revolution, evolution, twitter, Facebook, democratization, media technology, Arab Spring, French Revolution.

1 Introduction

It is indisputable that the recent developments in the Middle East are of global magnitude. The so-called "Arab Spring" got worldwide social, political and media attention due to its characteristics and connectedness with local, regional and international interests. From the starting point in Tunisia and spreading over to Egypt and other countries, people are revolting against suppressive forms of government and leadership. One of the biggest achievements consists in an emerging public dialogue within the societies and different social groups of the Middle East about the governing principles of everyday life. This has been possible through the ousting of long-term presidents in the region. Social media such as Facebook and twitter was mentioned by researchers and the press as an enabler for these happenings. Therefore, what happened in the Middle East is often referred to as "social media revolution" [1].

Nowadays, the use of social media is an established component for public and private communication in the region. The Egyptian use of these services increased through the January 25th revolution tremendously. It moved mainstream for a lot of people. The majority, however, is still not connected due to a high level of illiteracy and a lack of accessibility to information technology.

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 352–358, 2013.

Nevertheless, the importance of social media cannot be denied. Therefore, a lot of groups are debating its effects on revolutions. Some argue that the Internet is a double-edged sword that could be a potential tool for suppression [2]. Others clearly state that social media did have a significant role beyond being a source of news: "We use Facebook to schedule the protests, Twitter to coordinate, and YouTube to tell the world" [3].

The actual role of social media, however, whether supporting or causing the developments during the Arab Spring, can only be determined by putting the events into a broader context. Thus, the comparison with historic social developments on the one hand and technological advances on the other hand will be the focal point for the following chapters in order to see if the Arab Spring can be named a "social media revolution".

2 (Media) Revolutions

2.1 Revolutions

The concept of "revolution" is extremely diverse without a unified definition. Its Latin origin "revolution" means "turn around". This refers to "change" in the broadest sense. This change can be social, political, technical, cultural, psychological, personal or any other domain that would change fundamental principles. In the context of this paper, we look at revolutions of political and social systems. Tanter and Midlarsky [14] define four different types of revolutions based on characteristics such as mass participation, duration, domestic violence and intensions of the insurgents. Table 1 provides an overview of the types and their respective characteristics.

Table 1. Characteristics of revolutions according to Tanter and Midlarsky [14]

Type of revolution	Mass participation	Duration	Domestic violence	Intentions of insurgents
Mass revolution	High	Long	High	Fundamental changes in the structure of political authority and social system
Revolutionary coup	Low	Short to moderate	Low to moderate	Fundamental changes in the structure of political authority and possibly some change in the social system
Reform coup	Very low	Short, sometimes moderate	Low	Moderate changes in the structure of political authority
Palace revolution	None	Very short	Virtually none	Virtually no change

According to Tanter and Midlarsky [14], the French Revolution in 1789 as well as the German Revolution in 1848 could be categorized as mass revolutions. Revolutionary coups would be found in Turkey (1919), Germany (1933) and Egypt (1952). Argentina (1955), Syria (1956) and Pakistan (1958) experienced a reform coup while Venezuela and Brazil had a palace revolution in 1948 and 1955 respectively.

2.2 Media in Revolutions

Based on the schema for the categorization of revolutions mentioned above, the Arab Spring can be defined as a mass revolution. As such, it is clearly not an individual but a group phenomenon. Like every mass movement, communication between the different groups and actors plays a central role for the course of events. Due to its very nature, this mass communication is only possible if it is mediated in order to overcome the limitations of time and space. Therefore, it can be concluded that media technology tend to play an important role for mass revolutions in order to address these challenges.

Already the French Revolution "can be considered an epochal media event" [4]. The pre-industrialized printing press allowed the distribution of information in an amount and speed not known before. Tens of thousands of pamphlets were printed and distributed. The number of newspapers grew from just a couple "to over three hundred weekly and daily newspapers between 1789 and 1790" [4], while most of them were produced by individuals. This was possible only due to the low production costs. It is estimated that the total number of newspapers created during the revolution adds up to around 1,600 [4]. A lot of content was transmitted as well with printed pictures due to an illiteracy rate of more than 50% for men and more than 70% for women at the time [5].

In conclusion, media technology can be seen as an important component for the information and mobilization of the masses already in the 18th century. Further examples for this supporting role from recent history include the so-called twitter revolution in Iran [6] and the Multimedia or SMS revolt in the Philippines [7]. They show clearly that access as well as low cost of production and distribution account for possible easy dissemination of new ideas by individuals. Normally, television would not be considered as having these advantages. The protesters in Eastern Germany, however, were able to use the Western television and other media for informing the population of the former GDR on the protests that emerged increasingly in 1989 [8].

In summary, it can be stated that media technology, easy access and low entry barriers played always an important role in mass revolutions throughout history. It is not a unique consequence of the latest developments in information and communication technology. Therefore, the recent events in the Middle East might have happened sooner or later with a different course of events even without the use of social media.

2.3 Revolutionary Timelines

Looking at mass revolutions in the last centuries, it turns out that they are composed of single events and actions of different magnitudes. Although the Egyptian revolution is

referred to as the "January 25th Revolution", there is a set of previous events leading up to it [9]. The same can be observed for the time after the presidents of the region were ousted. The process of negotiating the new principle foundations of everyday life is contains violent events with a smaller magnitude than the main happening. This can be seen in both successful revolutions such as the French Revolution that peaked in 1789 and not successful ones such as the German Revolution in 1848.

Figure 1 shows a smoothed graphical representation of such a typical course of events. The horizontal axis is considered the timeline from left to right while the vertical axis refers to actions of opposing parties with different magnitudes. In the case of the Egyptian or Tunisian revolution, the central peak would relate to the ousting of the president. The waves on the left are the events and counter-events preceding the ousting. The waves after the peak on the right refer to events that represent the struggle of power between different parties. Interestingly, this course of events can be represented mathematically by the formula:

$$f(x) = \sin(x)/x$$

The following chapter is going to look at possibilities to describe different mass revolutions throughout the history by this formula.

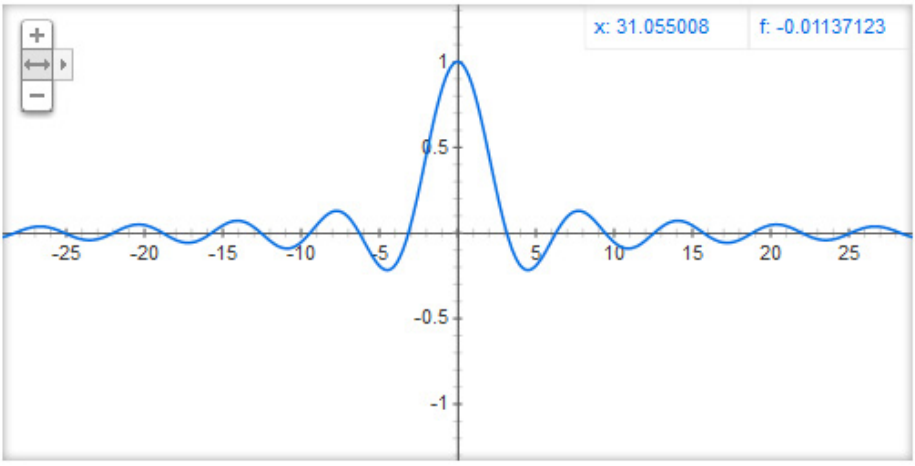


Fig. 1. Graphical representation of the course of revolutions [as displayed by google.com searching for f(x)=(sin(x)/x)]

3 The Changing Characteristic of Media Technology

The previous chapter concludes that media technology was always an important component of mass revolutions. So what has changed since the French Revolution? What are the differences between the technology available then and now?

It has been argued that mediated communication is essential for dissimilating information and ideas to the masses. Looking at the historical developments of other domains helps to understand the most important changing characteristic.

The sector of transportation saw tremendous changes in the last centuries. The daily distance possible to travel in the 19th century with a carriage drawn by horses was about forty miles maximum [10]. Today's high speed trains allow traveling this distance in 12min while it takes a plane about 4 minutes only. Besides the increased comfort, the speed of transportation is 165 times faster than centuries ago.

The changes that occurred in the business and financial sector are even more drastic. Before the invention of the telegraph, it took traders at the stock exchange between days and weeks in order to finish a transaction. The New York Stock Exchange today is able to close a transaction in 0.5ms [11]. The factor for the increase in speed is of hundreds of millions.

The distribution of news did undergo a similar development. According to [12] it took from 21 October 1805 to 6 November 1805 to publish the Battle of Trafalgar in the London Gazette. Similarly "the French invasion of Russia on 24 June 1812" was "published in The Times on 13 July 1812". Nowadays, a tweet or post on Facebook is published in a couple of seconds. The combination of influencers with a lot of connections and average users that "quickly convey the information from one neighbor to another" [13] results in the phenomenon of spreading the information faster than established media would do.

In summary, speed of distribution can be seen as the changing characteristic of technology. Taking both the changes in speed and the mathematical representation of revolutions presented in the previous chapter, it even might be possible to come up with a formal way of describing the course of revolutionary events by taking into account various factors such as speed of communication provided by type of media technology, number of participants and intentions of insurgents.

4 What's Next?

The diffusion of activities, statements and information increased tremendously through the usage of social networks such as Facebook and twitter. Compared to the high-speed trading executed at the stock markets, however, it is still in its infancy. The digitalization of the message with cameras and keyboards together with the distribution is still in the realm of seconds or even minutes. Projecting the latest technological advances in the fields of ubiquitous computing and augmented reality, however, could lead to a boost in the speed of diffusion that would enable types of revolutions not yet known.

A simple scenario combining the Google Glasses with an iWatch measuring and combining a huge set of data could lead to a true distribution of perceptions, emotions and happenings in real-time. Thus, if a huge amount of people share and synchronize such information, the network effect might be bigger and more influential than all the mass revolutions seen before.

5 Conclusion

This paper looked at the role of media technology in the recent social and political developments in the Middle East. The so-called "social media revolution" has been put into a broader technological and historical context in order to gain a deeper understanding of its impact. It has been shown that principal characteristics such as easy access and low entry barriers in terms of costs were already one of the main factors that support the developments in the French Revolution in 1789. The only basic element that changed constantly over the centuries was the speed of creation and transmission of information. Therefore, it is argued that the Arab Spring should rather be called a movement supported by media technology rather than a "social media revolution". The events would not have happened without the social and political reality on the ground. This reality, however, could have led to the same movement even without social media involved.

References

1. Tayler, K.: Arab Spring really was social media revolution. TG Daily (2011), http://www.tgdaily.com/software-features/58426-arab-spring-really-was-social-media-revolution (last access: March 1, 2013)
2. Shanthi, K., Taylor, B.: Open Networks, Closed Regimes: The Impact of the Internet on Authoritarian Rule. Carnegie Endowment (2003)
3. Chebib, N., Sohail, R.: The Reasons Social Media Contributed to the 2011 Egyptian Revolution. International Journal of Business Research and Management (2011)
4. Reichardt, R.: The French Revolution as a European Media Event (2012), http://www.ieg-ego.eu/en/threads/european-media/european-media-events/rolf-reichardt-the-french-revolution-as-a-european-media-event (last access: March 1, 2013)
5. Histoire de la scolarisation en France, http://fr.vikidia.org/wiki/Histoire_de_la_scolarisation_en_France (last access: March 1, 2013)
6. Burns, A., Eltham, B.: Twitter Free Iran: an Evaluation of Twitter's Role in Public Diplomacy and Information Operations in Iran's 2009 Election Crisis. In: Communications Policy & Research Forum 2009, November 19-20. University of Technology, Sydney (2009)
7. Coronel, S.: The media, the market and democracy: The case of the Philipines. Javnost=The Public 8(2), 109–124 (2001)
8. Friedensgebete und Montagsdmonstrationen Bundeszentrale für politische Bildung und Robert-Havemann-Gesellschaft e.V. (2008), http://www.jugendopposition.de/index.php?id=638 (last access: March 1, 2013)
9. Wolman, D.: The Digital Road to Egypt's Revolution. The New York Times (2012), http://www.nytimes.com/interactive/2012/02/12/opinion/sunday/20120212-tahir-timeline.html (last access: March 1, 2013)
10. Ritter, W.: Fremdenverkehr in Europa. Eine wirtschafts- und sozialgepgraphische Untersuchung über Reisen und Urlaubsaufenthalte der Beweohner Europas. Leiden (1966)

11. Schewe, P.: Relativistic trading: The speed of light isn't fast enough for some market transactions (2010), http://phys.org/news/2010-11-relativistic-isnt-fast-transactions.html (last access: March 1, 2013)
12. How fast did news travel? http://www.historyhouse.co.uk/articles/speed_of_news.html (last access: March 1, 2013)
13. Doerr, B., Fouz, M., Friedrich, T.: Why Rumors Spread Fast in Social Networks. Communications of the ACM 55, 70–75 (2012)
14. Tanter, R., Midlarsky, M.: A theory of revolution. The Journal of Conflict Resolution 11(3), 264–280 (1967)

Memotree: Using Online Social Networking to Strengthen Family Communication

Tsai-Hsuan Tsai[1], Yi-Lun Ho[1], Hsien-Tsung Chang[2], and Yu-Wen Li[2]

[1] Department of Industrial Design, Chang Gung University, Taoyuan, Taiwan
[2] Department of Computer Science and Information Engineering,
Chang Gung University, Taoyuan, Taiwan
ttsai.cgu@gmail.com

Abstract. This study describes the design and development of Memotree - a basic service platform for family social interaction. Memotree features a Family Tree which displays the family relationships and links among family members, with features to encourage family interaction based on the Family Communications Scale. The system and its associated hardware are designed to be optimally accessible to family members of different generations. Finally, Heuristic Evaluation was used to evaluate Memotree usability. Test subjects affirmed the Memotree concept and provided positive feedback.

Keywords: social networking sites, family tree, family communication scale, heuristic evaluation.

1 Introduction

In the context of filial piety, Asian societies place particular emphasis on kinship relations and rely on family trees to record the family's history and clan associations. Family trees are based on kinship ties, using the family surname to establish a family genealogy, recording the history of lineages and family factions descending from a common ancestor in a presentation integrating vertical and horizontal descent lines and spherical descent tables. The eldest in the family holds a place of particular respect within the family network[1, 2]. Traditional Chinese families typically consist of the elderly parents, adult children and non-adult grandchildren, commonly referred to as "three generations under one roof". In such arrangements, the grandparents are typically cast as the family historians, maintaining the family's history, recording the family's life experiences, and holding the family close together[3]. Generational interaction and mutual support not only contributes to general life satisfaction and well-being for the elderly, but also influences the family values and attitudes of the adult children and non-adult grandchildren. However, as society has become increasingly urban, industrial and commercial, an increasing number of families consist of conjugal or nuclear families, at the expense of extended families. These changes to residence style and family type have reduced the frequency of contact between generations which not only affects the psychological well-being of the elderly and the fulfillment of their social needs, but also raises critical issues for the family relations and support of modern families[4, 5].

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 359–367, 2013.

The application of internet technology today allows non-cohabitating family members stay in touch[6, 7]. The rapid development of modern online social networking sites (SNSs) has not only changed traditional social models, but have emerged as a critical social tool for today's youth[8]. Elderly people, despite their relatively low rates of use of computers or the internet, have been found to express a high degree of interest in learning to use such tools as a means of increasing their interaction with the younger generation[9]. Research has found that SNS-based communication do not replace face-to-face or telephone communication, but rather creates a new domain of family communication, providing family members with rich opportunities to maintain contact[10]. Unfortunately, currently most research into SNSs focus on personal networks rather than maintaining family relationships or the promotion of intergenerational exchange. Therefore, this study describes the design and development of Memotree - an only social platform suitable for promoting familial social interaction. Memotree not only imports the family genealogy concept from online SNSs, but also focuses on social interaction and communication between family members, thus promoting the maintenance of family relationships and intergenerational communication.

2 Memotree Design and Development

The product of collaboration between designers and engineers, Memotree offers social functionality similar to today's SNSs, but presented from the perspective of the user's family tree in the context of the user's family history and kinship structure, creating the elements of a Family Communication Scale to promote communication among family members. In addition, Memotree is designed to take into account the difference between older and younger people in terms of usage habits and experience to provide appropriate social platforms and hardware, ensuring that users of different generations are able to get the most out of the experience.

2.1 Importing the Family Tree Design Concept

In this research, the Family Tree feature is used to diagram the user's family relationships, family's lineage and kinship organization. In Memotree, the user can manually manage the Family Tree by adding or deleting family members, or using the Calendar feature to record family holidays, birthdays, anniversaries, etc. It's worth mentioning that, in the Family Tree network, the relationships are closest among the three above-mentioned generations of family members, thus Memotree prominently features these intergenerational relationships in its main display, allowing the user to locate themselves within the context of their preceding or succeeding generations (Fig. 1). Once the Family Tree is populated, it can automatically determine the kinship between various users. If a relationship meets the criteria for kinship, users can establish the relationship, eventually recreating their extended family. In addition, the system automatically recognizes the identity of the user within the family and the specific relationship with other members.

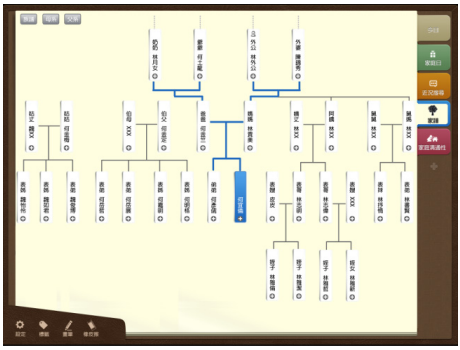


Fig. 1. Family tree

Fig. 2. MemoTree main page

2.2 Functional Design Based on the Family Communication Scale

To promote interfamily communication and interaction, this study based the design of Memotree functionality on Olson's Family Communication Scale (FCS), which identifies five general types of family communication skills by which the degree of family communication can be assessed: listening skill, speaking skill, self-disclosure, tracking, and respect and regard[11, 12]. This study first analyzed the characteristics of the social functions of current SNSs. Ten pairs of grandparents and grandchildren were then recruited as respondents, including 7 grandmothers and 3 grandmothers (with an average age of 72.4 years), along with 2 grandsons and 8 granddaughters (with an average age of 25.5 years). By evaluating the degree of family communication among these pairs and interviews regarding their family interaction needs, this study determined the gap between these needs and the functionality of current SNSs. Finally, based on the research results, we designed interactive functions to match the FCS results and users' family interaction needs (see Fig. 2). The interactive functions are described as follows:

- Listening skill: Listening skill is defined as empathy and attentive listening. Memotree's Listening functions include video conferencing and chat rooms which not only allow users to hear each other's voices but also help family members share opinions and viewpoints. However, video conferencing is generally easier for elderly users than text-based chat rooms.
- Speaking skill: Speaking skill is defined as talking to another person. Memotree's Speaking features include video conferencing and chat rooms which not only allow users to converse with others about personal issues but also help people express their feelings. In addition, the video feature allows users to sense each other's presence while speaking, thus mimicking face-to-face interaction.
- Self-disclosure: Self-disclosure is defined as sharing one's feelings with others. Memotree's Self-disclosure features center around status updates and photo sharing. The status update feature allows users to share their current mood or recent events with family members. In addition to allow test subjects to freely

share personal information with privacy, Memotree's Self-disclosure space is divided into public and private areas, providing users with the ability to share feelings and thoughts so that family members might better understand one another.

- Tracking: Tracking is defined as pausing on a topic. Memotree offers tracking through its comment feature. Family members can post text-based or icon-based replies on a given topic, thus providing family members emotional support and encouragement. It's worth pointing out that the Memotree system collects messages for each topic and, based on the message properties and the topic header, saves them to the Family Calendar system, allowing all family members view the daily, weekly and monthly activity of other family members, thus giving them insight into their recent situation.
- Respect and Regard: Respect and Regard is defined as the emotional side of family communication. Memotree includes a Family Calendar system which records important family events such as birthdays, memorial days and anniversaries. Memotree users can specify certain days as Family Days, setting the Family Calendar to send out appropriate messages to other family members to show respect and concern.
- Family Communication Measure: Taking Olson's Family Communication Scale as the basis for evaluation, the Memotree system collects statistics on monthly activity by family members to calculate the level of family communication. In addition, the system will also provides statistics on average number of communication instances between family members which not only helps users better understand the level of interfamily communication, but also provides timely recommendations and reminders to help users and their families maintain better contact (Fig. 3).

Fig. 3. Family communication measure

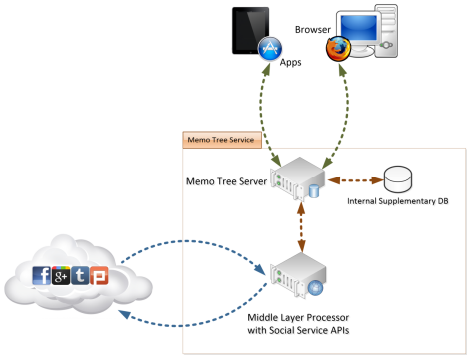


Fig. 4. System architecture of Memotree Service

2.3 System Interface and Hardware Designed to Accommodate User Habits

The Memotree user interface provides a Family Notepad to emphasize interaction between family members and to record family events (Fig. 2). The Memotree system also links to popular SNSs such as Facebook and Google+, allowing younger users to

simultaneously interact with the Memotree system and other social platforms. In terms of hardware, the Memotree system supports most browsers and smartphone devices, including tablets, thus allowing elderly users who are potentially unfamiliar with PC operation to easily operate the Memotree family communication platform. Figure 4 illustrate the overall Memotree system architecture. The system is divided into a server side and a client side. The server side includes an intermediary server responsible for managing all APIs for community service software and communications. The intermediary server is designed so that updating the server software obviates the need for API updates and updates to the client side software. In addition, the intermediate server also conducts data Fusion, synchronously a user's data throughout different community services. The Memotree server is responsible for all system operations and controls. Most data is obtained through the server, which synchronously updates user information with community services. The Memotree system only records some statistical data and account data to an Internal Supplementary DB. Users can access their Memotree services through a tablet-based application or through their PC browser. Overall, the Memotree system uses Fusion to integrate the band-end community services and uses the Re-Render concept to re-interpret community services.

3 Validation

Heuristic Evaluation was proposed by Nielson as a means of evaluating usability, and is now widely used to assess a variety of systems[13, 14]. For example, Hart et al. applied the approach to evaluate interface usability problems on Facebook[15], Pinelle et al. applied it to assess interactive games[16], and Lee & Kozar used it to evaluate the relationship of usability to purchase intention on e-commerce sites[17]. Accordingly, this study uses Heuristic Evaluation to assess usability of the Memotree family social interaction platform.

3.1 Test Subjects and Experimental Design

We invited three experts in user interface design or ergonomic engineering to participate in the usability assessment of the Memotree system, including two experts with rich experience in information engineering system design, and one expert with a strong background in interface design and usability. The experts were asked to assume the role of a general user for system evaluation. All interaction with the system was observed and recorded to identify problems, solicit opinions, determine usability design flaws and analyze their causes. In addition, to obtain the point of view of actual users, the solicited 42 subjects between the ages of 20 to 30 to test the Memotree system. After testing, subjects were asked to provide recommendations regarding the system interface and functionality.

3.2 Results

We invited three experts in user interface design or ergonomic engineering to participate in the usability assessment of the Memotree system, including two experts with rich experience in information engineering system design, and one expert with a strong background in interface design and usability. The experts were asked to assume the role of a general user for system evaluation. All interaction with the system was observed and recorded to identify problems, solicit opinions, determine usability design flaws and analyze their causes. In addition, to obtain the point of view of actual users, the solicited 42 subjects between the ages of 20 to 30 to test the Memotree system. After testing, subjects were asked to provide recommendations regarding the system interface and functionality.

(a) Expert evaluation results

Heuristic Evaluation uses ten usability evaluation principles to identify potential interface usability issues[13, 14]. The three experts were asked to assess the Memotree system according to the ten criteria in Neilsen's Heuristic Evaluation. In the inspection process, the researchers recorded whether the experts considered the system to present usability issues for each criteria, with an increasing number of reports indicating an increased need for remedial action. The expert evaluation results for the ten Heuristic Evaluation criteria are as follows:

- Visibility of system status: In terms of system status presentation, the experts indicated that the user could easily recognize his or her location within the system, and that system information displays were clear and useful. However, they felt that the display text size may be too small for older users, and thus cause them difficulty in reading.
- Match between system and the real world: The experts felt that the Memotree system corresponds well to user operation, and the user can be made immediately aware of the changes caused by his/her actions within the system.
- User control and freedom: In terms of user control and freedom, the Memotree system is fairly uncomplicated. Action selection are clearly differentiated, and the experts felt that further user control mechanisms were unnecessary.
- Consistency and standards: The experts felt that the Memotree system's color scheme and type face selections fulfilled design principles for consistency, but some minor adjustments should be made to ensure the consistency of some icons.
- Error prevention: Memotree system does not provide an error-handling mechanism. The experts felt that the system's relatively uncomplicated design and ease of use should reduce the user error rate.
- Recognition rather than recall: The experts felt that the Memotree system's menu system was highly navigable. This reduces the amount of information the user needs to remember, thus reducing the user's cognitive loading. However, the expert evaluators also mentioned that the clickable buttons should be adjusted to clearly differentiate clickable links and general information.
- Flexibility and efficiency of use: The three experts noted that all Memotree system functions are shown on the system interface, such that users do not need to search

for the location of function. Therefore, although the system does not provide any shortcuts, the expert evaluators all affirmed the current approach in terms of flexibility and efficiency of use, and did not raise any further recommendations.

- Aesthetic and minimalist design: The expert evaluators felt that the Memotree system interface messaging was concise, jargon-free, and easy to understand. In addition, the system page only displays information needed to help users easily recognize and operate the desired function.
- Help and documentation: Memotree system does not provide user manuals and guides. The three experts recommended including simple instructions to help new users quickly learn to operate the Memotree system.

(b) User viewpoint

Forty-two young test subjects, all familiar with the operation and function of online SNSs (e.g., Facebook, etc.) were invited to help test the Memotree system. To solicit their viewpoint, the testers were first asked to use the system and then provide comments and suggestions for the system interface and functionality. Feedback from the test subjects generally held that the interface design was attractive, the functions are well-defined, and the system provides meaningful and effective information. Three functions were singled out for specific recognition and appreciation: (1) the Family Tree specifies the kinship relationship between various family members as they interact in the system, (2) the interface design of the Family Calendar is not found in other current SNSs, and (3) the Family Day function allows family members to express emotion and concern, thus promoting cohesion among family members. Selected tester responses are organized below:

- Family Tree: Unlike other current SNSs, the Memotree system only provides communication between family members. Test subjects felt that the Memotree system's Family Tree feature depicts the identity of individual family members and their specific relationships with other members, while also providing the basic information and current status of each member (e.g., birthday, recent photos and current status). This allows users to get to know their family members better and also helps promote the continuation and maintenance of the traditional Chinese sense of family and familial culture. In addition, the Family Tree function automatically searches for family relationships and contacts between various members of an extended family, thus creating an elaborate kinship diagram for an extended family.

  User quote: "… This family tree feature is amazing – it organized my entire family's information…"

  User quote: "… I think this automatic editing feature in the family tree is very good – it can let me clearly see who's in my family…"

- Family Calendar interface design: Designed on the concept of conventional calendars, the Family Calendar interface is easy for novice users (especially elderly users) to intuit and learn, thus reducing the amount of time and effort needed to learn to use the system. The 42 young test subjects indicated that the interface

design was attractive, providing a simple and intuitive mode of operation which meets the cultural expectations and habits of elderly users, such as the inclusion of the lunar calendar. In addition, the calendar offers links to day-to-day family events and records the feelings of family members.

> *User quote: "… I'd show my grandfather the lunar calendar feature – he'll like that."*
>
> *User quote: "… the calendar is organized on a daily basis, allowing me to clearly see how everyone in my family is doing…"*
>
> *User quote: "… it's like reading a diary. It's very convenient and provides a way for family members to interact."*

- Promotion of emotional expression and concern through Family Day: The Family Calendar not only allows users to set specific Family Days (e.g., birthdays, anniversaries, family activity days, etc.), but the system also reminds users of upcoming Family Days to they can express their concern for the family. From an emotional point of view, this kind of interaction is a type of rich emotional transfer which can increase the intimacy between family members. In addition, the test subjects indicated that the inclusion of icons and emoticons along with text gave them a direct and convenient way to express their emotions and concern.

> *User quote: "On birthdays, I want to send my family members flowers or a birthday cake, and these cute icons help me express my best wishes…"*
>
> *User quote: "… these little gift [icons] are really great. I hope we can have a bigger selection of gifts. It would make me very happy to received such gifts…."*

4 Conclusion

To extend traditional Chinese family values and ancestral culture, and to adapt these concepts to modern popular digital social interaction modes, this study developed Memotree – a basic service platform for family social interaction. Memotree features a Family Tree which displays the family relationships and links among family members, with features to encourage family interaction based on the Family Communications Scale. The system and its associated hardware are designed to be optimally accessible to family members of different generations. Finally, Heuristic Evaluation was used to evaluate Memotree usability. Test subjects affirmed the Memotree concept and provided positive feedback. Specifically, the Family Tree, Family Calendar and Family Day features were seen as being conducive to promoting strong family bonds and rich communication and interaction among family members, thus satisfying the usage needs of different generations of the same family.

Acknowledgments. The authors would like to thank the National Science Council, Taiwan, R.O.C., for supporting this research under grant NSC-99-2410-H-182-032.

References

1. FamilySearch (December 25, 2011), https://familysearch.org/
2. Ancestry.com. (December 25, 2011), http://www.ancestry.com/
3. Hagestad, G.O.: Transfers between grandparents and grandchildren: The importance of taking a three-generation perspective. Zeitschrift für Familienforschung, pp. 315–332 (2006)
4. Silverstein, M., Cong, Z., Li, S.: Intergenerational transfers and living arrangements of older people in rural China: Consequences for psychological well-being. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences 61, S256–S266 (2006)
5. Domingo, L.J., Asis, M.M.B.: Living arrangements and the flow of support between generations in the Philippines. Journal of Cross-Cultural Gerontology 10, 21–51 (1995)
6. Selwyn, N.: The information aged: A qualitative study of older adults' use of information and communications technology. Journal of Aging Studies 18, 369–384 (2004)
7. Lehtinen, V., Näsänen, J., Sarvas, R.: A little silly and empty-headed: older adults' understandings of social networking sites. Presented at the Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology, Cambridge, United Kingdom (2009)
8. Boyd, D.: Why Youth Social Network Sites: The Role of Networked Publics in Teenage Social Life. In: John, D., Catherine, T. (eds.) Foundation Series on Digital Media and Learning (2007)
9. Tsai, T.H., Ho, Y.L., Tseng, K.: An investigation into the social network between three generations in a household: bridging the interrogational gaps between the senior and the youth. Online Communities and Social Computing, pp. 277–286 (2011)
10. Chen, Y., Wen, J., Xie, B.: I communicate with my children in the game: Mediated Intergenerational Family Relationships through a Social Networking Game. The Journal of Community Informatics 8 (2012)
11. Barnes, H.L., Olson, D.H.: Parent-adolescent communication and the circumplex model. Child Development 56, 438–447 (1985)
12. Olson, D.H.: Circumplex model of marital and family sytems. Journal of Family Therapy 22, 144–167 (2000)
13. Nielsen, J., Molich, R.: Heuristic evaluation of user interfaces, 249–256 (1990)
14. Nielsen, J.: Enhancing the explanatory power of usability heuristics. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Celebrating Interdependence, pp. 152–158 (1994)
15. Hart, J., Ridley, C., Taher, F., Sas, C., Dix, A.: Exploring the facebook experience: a new approach to usability. In: Proceedings of the 5th Nordic Conference on Human-Computer Interaction: Building Bridges, pp. 471–474 (2008)
16. Pinelle, D., Wong, N., Stach, T.: Heuristic evaluation for games: usability principles for video game design. In: Proceeding of the Twenty-Sixth Annual SIGCHI Conference on Human Factors in Computing Systems, pp. 1453–1462 (2008)
17. Lee, Y., Kozar, K.A.: Understanding of website usability: Specifying and measuring constructs and their relationships. Decision Support Systems (2011)

Part IV

HCI for Business and Innovation

It Was Nice with the Brick So Now I'll Click: The Effects of Offline and Online Experience, Perceived Benefits, and Trust on Dutch Consumers' Online Repeat Purchase Intention

Ardion Beldad and Mariël Segers

Department of Corporate and Marketing Communication,
University of Twente, The Netherlands
a.d.beldad@utwente, m.segers@student.utwente.nl

Abstract. The present study investigated the factors influencing Dutch consumers' intention to continue purchasing from a brick-and-click clothes shop's online channel after an initial commercial exchange. Results on the online survey with 513 respondents reveal that their repeat online purchase intention is influenced by the quality of their previous online purchase experience with the retailer, their perception of the benefits associated with online purchasing, and their trust in the retailer. Analysis also shows that respondents' previous offline purchase experience and online purchase experience contribute to their trust in the retailer. Additionally, respondents' previous online purchase experience positively influences their perception of the benefits associated with purchasing online. The modified model also shows that respondents' trust in the retailer positively influences their perceptions of online shopping benefits.

Keywords: offline purchase experience, online purchase experience, online trust, repeat online purchase intention, blick-and-click.

1 Introduction

The increasing popularity of online shops has prompted an increase in the number of companies, which operate exclusively offline, to offer their products and services online. Apparently, shops need to sell their products and services through both offline and online channels not only to survive a stiff competition but also to better serve customers whose buying behaviors have become increasingly mediated by electronic technology, eventually resulting in the metamorphosis of online transactions from pure clicks to brick-and-clicks [19]. With this phenomenon comes significant attention to the relationship between a customer's interaction with a retailer in an offline environment and the customer's propensity to transact with that same retailer in the online environment.

One may assume that a company that operates purely online is at a disadvantaged position compared to a company with both offline and online channels, consequently referred to as brick-and-click companies, since companies with an offline presence

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 371–380, 2013.

are more likely to be trusted in the online environment than those without offline presence [16]. Online trust, as several studies have shown, is a strong determinant of people's intention to engage in exchanges with organizations online. Although trust in a company could positively influence customers' intention to purchase something from that company's online channel, other factors that could potentially increase such an intention deserved to be identified. What is clear is that the factors influencing the decision to buy online for the first time could be partly different from those that might increase customers' propensity to continue buying from a shop that operates exclusively online and from an online shop with an offline presence.

While empirical studies that have identified the determinants of first-time online purchase and repeat online purchase abound, specifically in the context of companies that sell exclusively through their websites, there is still a dearth of research into the factors influencing customers' willingness to purchase again from an online shop with an offline presence. This research primarily aims at addressing this gap. An online survey with a brick-and-click clothing shop in the eastern region of the Netherlands was conducted to test the research hypotheses.

2 Determinants of Online Repeat Purchase Intention

Companies continuously strive not only to attract customers to buy something from them but also to maintain customers' loyalty, primarily expressed in their propensity to engage in a repeat purchase. One question deserving research attention, however, pertains to the determinants of people's online repeat purchase intention, referring to the likelihood that people will buy something from the same company online after a first-time purchase.

2.1 Experience with the Shop both Offline and Online

People's decision to engage in an initial online transaction with a company is often predicated on factors such as trust, organizational reputation, and the expected benefits that can be derived from such form of transaction. However, upon the acquisition of transactional experience and experience with the transaction partner, people can already base their repeat purchase intention on first-hand information. While customers of shops that operate exclusively online could base their decision to buy again from the shop solely on their online purchase intention, those who have purchased something from the shop's offline and online channels can predicate their online repeat purchase intention on their transactional experience with both channels.

An important point to consider, however, is that experience does not entirely suffice to increase repeat purchase intention. As experience can be in a pole with satisfaction and enjoyment at one side and dissatisfaction to disappointment at the other side [2], a satisfying experience could logically increase repeat purchase intention. Indeed, studies have shown that a satisfying online transaction experience is a strong indicator of the aforementioned intention [22, 25].

It is also argued that people's experience with a shop offline is often translated into their buying behavior online, as the service obtained from a shop through its physical

outlet could positively impact customers' willingness to purchase from the shop through its online channel [24]. The following hypotheses, therefore, are advanced.

H1a : Customers' positive experience with a shop in the offline environment positively influences their intention to buy again from the shop's online channel.
H1b : Customers' positive experience with a shop in the online environment positively influences their intention to buy again from the shop's online channel.

A positive relationship between a positive experience and trust in an online shop exists as customers who are satisfied with their initial online exchange experience tend to be more trusting of the company in the online environment [5, 10, 18]. Based on this assertion, the next set of hypotheses is advanced.

H1c : Customers' positive experience with a shop in the offline environment positively influences their trust in the shop's online channel.
H1d : Customers' positive experience with a shop in the online environment positively influences their trust in the shop's online channel.

2.2 Trust in the Shop's Online Channel

Referred to as one party's reliance on the voluntarily assumed responsibility of another party to safeguard the rights and interest of the former [11], trust has been regarded crucial in prompting individuals and organizations to engage in various forms of exchanges with others. The success of online exchanges (e.g. e-commerce, e-government) is predicated on trust [7] considering the uncertainties and risks associated with online transactions characterized as distant, anonymous, and impersonal [18]. Certainly, trust does not eliminate risks although it is highly instrumental in reducing the degree and the type of risks (e.g. financial losses, information privacy violation) [3] people perceived from engaging in transactions with the trust (or distrust) target [3, 13, 23]

Initial encounters, when both parties have no prior knowledge about each other's performance, are partly predicated on initial trust, which is based on the knowledge about the trustee a trustor obtains from third parties [9] and on the expectation that the trustee will not defect for fear of negative outcomes [17]. After an initial encounter, the two parties' decision to engage in succeeding exchanges could already be grounded on trust resulting from the transacting parties' knowledge about each other [17]. Trust, therefore, matters not only in initial encounters but also in succeeding encounters. This prompts the next hypothesis.

H2 : Customers' trust in the (brick-and-click) shop's online channel positively influences their online repeat purchase intention.

2.3 Benefits of Online Shopping

The acceptance of novel technologies and newer ways of doings things is partly dependent on their usefulness, as the Technology Acceptance Model postulates [8]. The usefulness of online shopping can be measured in terms of the benefits it extends to customers such as convenience and comfort [6, 14]. The expected benefits people

associate with online shopping have been found to strongly influence their intention to buy something online for the first time [14].

As people gain more online transaction experience, nonetheless, the extent to which they were satisfied or displeased with their previous transactions could even serve as a relatively objective indicator of their evaluation of the benefits associated with an online transaction. In fact, empirical studies have shown that beliefs in the usefulness of computer-mediated commercial exchanges are an important determinant of online repurchase intention [25]. This prompts the hypothesis below.

H3 : Customers' perception of the benefits of purchasing online positively influences their intention to continue purchasing from a (brick-and-click) shop's online channel.

3 Method

3.1 Sampling

An online survey was conducted with consumers of one brick-and-click clothes shop in the eastern part of the Netherlands. The retailer, which sells primarily women and children apparel, also assisted the researchers in distributing the questionnaires to the study's target respondents since the retailer maintains a database of contact information of consumers who have purchased a product from both the offline and online channels of the retailer. With the assistance of the retailer, a link to the online questionnaire was sent to 8,147 customers whose e-mail addresses were stored in the retailer's electronic database at the time of the study. The focus of the study required that only respondents who have purchased from the retailer through both its offline and online channels should be included in the study. After a four-week data collection, 706 filled out online questionnaires were returned. However, 193 questionnaires had to be removed since they were not completed, resulting in 513 questionnaires used for analysis or a final response rate of 6.29%.

3.2 Respondents

As the retailer used for this study sells women and children apparel, almost all respondents (N=499, 97%) were women. Respondents' age ranged between 21 and 60 (M = 39.06, SD =7.99). In terms of level of education, 57% (N = 291) of the respondents have obtained higher education (university or bachelor degrees). Most respondents indicated to have high levels of Internet experience, as those who have been using the Internet for more than 11 years accounted for 45% (N = 233) of the study's sample.

3.3 Research Instrument

New items were formulated for the constructs 'offline experience with the shop' and 'online experience with the shop' – with three items for each construct. Examples of items for 'offline experience' included 'My experience with the shop offline was

positive' and 'My previous purchase with the shop offline proceeded without problems'. For 'online experience', typical items included 'My experience with the shop online was positive' and 'I never had negative experiences with the shop's online channel'. Cronbach's alpha scores for 'offline experience' and 'online experience' were pegged at .90 and .93, respectively.

The three items to measure 'trust' (α = .85) were based on the scales of Carter and Belanger [4] and Jarvenpaa et al. [13]. Examples of items included 'I trust the shop's online channel' and 'I believe the shop's online channel is trustworthy'. 'Perceived benefits of online shopping' (α = .85) were measured with four items by Kim, Ferrin, and Rao [14]. Examples of items for the construct included 'Buying from the shop's online channel saves me time' and 'I can always buy something from the shop's online channel anytime I want to'. The dependent variable 'repeat online purchase intention' (α = .90) was measured with three newly formulated items such as 'I would most likely buy an item from the shop's online channel again' and 'I am intending to purchase an item from the shop's online channel anytime soon'.

Presented on Table 1 are the Cronbach's alpha, mean, and standard deviation scores for the different constructs of the study. All items used for the different constructs were measured on a five-point Likert scale (with 1 corresponding to 'fully agree', 2 'agree', 3 'neither agree nor disagree, 4 'agree', and 5 'fully disagree'.

Table 1. Cronbach's α, mean, and standard deviation scores for the different constructs

Constructs	Cronbach's α	Mean	Std. Deviation
Offline experience with the shop	.90	1.55	.71
Online experience with the shop	.93	1.92	.84
Trust in the shop's online channel	.85	1.80	.56
Perceived benefits of online shopping	.85	2.05	.66
Repeat purchase intention	.90	2.34	.71

4 Results

To test the different research hypotheses and to determine whether the proposed research model fits the data, structural equation modeling (SEM) using AMOS 18.0 was performed. The use of SEM subscribed to the two-step approach proposed by Anderson and Gerbing [1], which indicates that the measurement model should be assessed first prior to testing of the structural model. Model fit was assessed using the following indices: RMSEA (root mean square error of approximation), CFI (comparative fit index), TLI (Tucker-Lewis index), and SRMR (standardized root mean square residual) [15, 20]. CFI and TLI values should be greater than .95 [12], the RMSEA value lower than 0.07 [21], and the SRMR value below 0.08 to reach a conclusion that a relatively good fit between the proposed model and the observed data exists [12, 15].

Results of Confirmatory Factor Analysis (CFA) show that the measurement model excellently fits with the data [$X^2(80) = 265.37$, $X^2/df = 3.32$, p = .00, CFI = .97, TLI = .96, SRMR = .03, RMSEA = .07 (CI: .06, .08)]. Test of the structural model resulted in a relatively acceptable fit [$X^2(83) = 313.76$, $X^2/df = 3.78$, p = .00, CFI = .96, TLI = .95, SRMR = .05, RMSEA = .07 (CI: .07, .08)]. However, inspection of the

modification indices suggested that model fit could still be improved by establishing a causal relationship between two independent variables. The original model, therefore, was modified by adding a path from 'online experience' to 'perceived benefits' and from 'trust' to 'perceived benefits'. The test of the modified model yielded a slightly improved fit [$X^2(82) = 278.28$, $X^2/df = 3.39$, $p = .00$, CFI = .96, TLI = .95, SRMR = .04, RMSEA = .07 (CI: .06, .08)].

Analysis shows that respondents' repeat online purchase intention is influenced by the quality of their previous online purchase experience with the retailer ($\beta = .32$), their trust in the (brick-and-click) shop's online channel ($\beta = .13$), and their perception of the benefits associated with purchasing online ($\beta = .39$). This results to the acceptance of hypotheses 1b, 2, and 3. However, the quality of respondents' previous offline purchase experience with the retailer has been found to have no influence on repeat online purchase intention, leading to the rejection of hypothesis 1a.

Moreover, analysis indicates that respondents' previous offline purchase experience ($\beta = .34$) and online purchase experience ($\beta = .46$) contribute to their trust in the shop's online channel. Thus, hypotheses 1c and 1d are accepted. The modified model also shows that respondents' previous online purchase experience positively influences ($\beta = .36$) influences the perceived benefits of purchasing online. Furthermore, respondents' trust in the (brick-and-click) shop's online channel positively influences perceptions of the benefits of online shopping ($\beta = .31$)

5 Discussion

For retailers that used to operate exclusively offline, the expansion of product and service delivery to the online environment entails a thorough understanding of the ways to ensure that customers who purchase offline would also exchange in a commercial exchange with the retailer through its online channel. There certainly are copious studies into the factors influencing initial online purchase intention and repeat online purchase intention, especially in relation to shops with online channels only. However, adequate academic attention has not been extended to the mechanism behind people's willingness to purchase through the online channel of a retailer with an offline presence.

Results of this study show that people's intention to purchase something again from a brick-and-click retailer's online channel is predicated on three important factors, namely, the quality of their previous online commercial exchange, their level of trust in the retailer's online channel, and the perceived benefits of online shopping. These findings strongly support the results of previous empirical studies into the impact of these three factors on purchase and repeat purchase intentions.

It should be noted that of the three, however, the perceived benefits of online shopping strongly determine the extent to which customers would opt to purchase an item from an offline retailer's online channel after an initial exchange. While this variable lies outside a specific retailer's influence, the retailer can still ensure the continuance of an online purchase among its customers by providing them with a satisfactory online transaction experience.

Although it was initially surmised that the quality of customers' experience with a retailer offline would influence their repeat online purchase intention, the study's results indicated that the former has no impact on the latter. This is surprising since one would expect that the extent to which customers were pleased with their purchase through the retailer's offline channel would determine their willingness to shop through that retailer's online channel. A possible explanation for this is that when deciding whether or not to shop online, customers may not really pay attention to their offline exchange experience with the retailer behind the online channel.

Even if the evaluation target – for quality of previous experience and trust – is the same (the retailer), the channels used for the commercial exchange differ in several aspects. Purchasing an item offline provides a customer with the possibilities to interact with a flesh-and-blood salesperson and to feel or try a product, which would totally be impossible if one opts to buy a similar item through the retailer's online channel. Thus, when deciding whether or not to continue buying from a retailer's online shop, customers might pay more attention to the quality of their previous online transaction than to their purchase experience offline.

The fact that trust in the retailer's online channel positively influences customers' repeat purchase intention amplifies the need to understand how trust could be developed. Trust has been known to evolve depending on the depth of the relationship between a trustee (in this case, the retailer) and a trustor (the customer). While initial encounters could be predicated on trust that emerges from the knowledge a trustor obtains from others, succeeding encounters could already be based on trust that develops out of the quality of the interaction between the two parties. Studies have shown that customers who were satisfied with their previous transactions with a company are more likely to continue trusting the company after an initial encounter.

Results of this research strongly supports the relationship between quality of previous transaction experience and trust online. The quality of respondents' previous online exchange experience with a blick-and-click is a strong determinant of their trust in the retailer's online channel. Although the quality of respondents' offline exchange experience does not impact repeat online purchase intention, results further reveal that it is an important determinant of trust in the retailer's online channel. While trust in the shop, in general, is not identified as a construct in this study, this may have been captured by the quality of the respondents' offline exchange. Customers who were satisfied with their transaction with the retailer in the offline environment have a strong reason to trust the retailer, in general, and such trust could influence the extent to which they would trust the retailer's online channel.

Modification of the structural model indicates that factors such as 'quality of previous online experience' and 'trust in the retailer's online channel' positively influence respondents' perception of the benefits of online shopping. An important implication of this finding is that online customers' estimation of the benefits of online shopping is partly predicated on the quality of their previous online transaction experience. And this further suggests that when the previous online transaction with the retailer was not gratifying, customers would be most likely inclined to undermine the merits of online shopping.

Additionally, as the analysis shows, the benefits of online shopping would hardly be appreciated if the channel used for the online transaction is deemed untrustworthy. Trust, therefore, is a critical factor as it substantially impacts not only online customers' behavioral intention but also their attitude towards the object of the intention – online shopping.

6 Implications

Results of this study have important implications not only for electronic commerce researchers but also for brick-and-click retailers. While it is known that customers' experience with a retailer offline does not prompt repeat online purchase intention, such an experience, which should primarily be satisfactory, is an important determinant of their trust in the channel the retailer uses for commercial transactions online. Brick-and-click retailers, therefore, should continuously strive to provide their customers with a gratifying experience offline as it can influence customers' trust in the retailer in the online environment, which partly determines their repeat online purchase intention and the extent to which online shopping is seen as beneficial.

What is more important is that brick-and-click retailers should incessantly ensure that online shoppers are satisfied with their online shopping experience, as it is a critical determinant of repeat online purchase intention. Furthermore, such an experience, as this study's results reveal, shapes people's evaluation of the benefits of online shopping, which positively influences repeat online purchase intention. Although it is beyond a brick-and-click retailer's ability to influence people's belief in the benefits of online shopping, retailers could still have an impact on it by maintaining customers' trust in the channel for an online exchange and by guaranteeing that customers would be satisfied with such transaction – the two pivotal determinants of the extent to which online shopping is viewed as advantageous.

The model tested in this study is remotely comprehensive. The effect of customers' experience with a brick-and-click retailer on trust in the retailers' online channel is interpreted as a consequence of the trust people have in the retailer, which results from customers' positive exchange experience with the retailer. However, trust in the retailer, in general, proves to be an important variable that needs to be included in a more comprehensive model for repeat online purchase intention.

Results of the current study might have been skewed by the unusually high number of female survey respondents, considering the type of brick-and-click retailer used for the research. One can, therefore, assume that the findings could be a reflection of the mechanism behind female shoppers' repeat online purchase intention. The relatively high impact of perceived benefits of online shopping on the behavioral intention of interest might be a consequence of having a large proportion of women in the sample, as it can be assumed that female shoppers might consider online shopping benefits such as convenience and time-saving as decisive factors for their repeat online purchase intention. The model, therefore, could also be tested with male respondents to see whether or not the factors influencing repeat online purchase intention vary according to the gender of shoppers.

References

1. Anderson, J.C., Gerbing, D.W.: Structural equation modeling in practice: A review and recommended two-step approach. Psychological Bulletin 103, 411–423 (1988)
2. Beldad, A., De Jong, M., Steehouder, M.: How shall I trust the faceless and the intangible? A literature review on the antecedents of online trust. Computers in Human Behavior 26, 857–869 (2010)
3. Beldad, A., Van Der Geest, T., De Jong, M., Steehouder, M.: Shall I tell you where I live and who I am? Factors influencing the behavioral intention to disclose personal data for online government transactions. International Journal of Human-Computer Interaction 28(3), 163–177 (2012)
4. Carter, L., Belanger, F.: The utilization of e-government services: Citizen trust, innovation, acceptance factors. Information Systems Journal 15(1), 5–25 (2005)
5. Casalo, L.V., Flavian, C., Guinaliu, M.: The influence of satisfaction, perceived reputation and trust on a consumer's commitment to a website. Journal of Marketing Communications 13(1), 1–17 (2007)
6. Chiu, C.M., Wang, E.T.G., Fang, Y.H., Huang, H.Y.: Understanding customers' repeat purchase intentions in B2C e-commerce: the roles of utilitarian value, hedonic value and perceived risk. Information Systems Journal (2012), doi:10.1111/j.1365-2575.2012.00407.j_407
7. Corritore, C.L., Kracher, B., Wiedenbeck, S.: Online trust: Concepts,evolving themes, a model. International Journal of Human-Computer Studies 58, 737–758 (2003)
8. Davis, F.D.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 13(3), 319–340 (1989)
9. Doney, P.M., Cannon, J.P., Mullen, M.R.: Understanding the influence of national culture on the development of trust. Academy of Management Review 23(3), 601–620 (1998)
10. Flavian, C., Guinaliu, M., Gurrea, R.: The role played by perceived usability, satisfaction, and consumer trust on website loyalty. Information & Management 43, 1–14 (2006)
11. Hosmer, L.T.: Trust: The connecting link between organizational theory and philosophical ethics. Academy of Management Review 20(2), 379–403 (1995)
12. Hu, L., Bentler, P.M.: Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling 6, 1–55 (1999)
13. Jarvenpaa, S.L., Tractinsky, N.: Consumer trust in an Internet store: A cross-cultural validation. Journal of Computer-Mediated Communication 5(2) (1999), `http://jcmc.indiana.edu/vol5/issue2/jarvenpaa.html` (retrieved)
14. Kim, D.J., Ferrin, D.L., Rao, H.R.: A trust-based consumer decision-making model in electronic commerce: The role of trust, perceived risk, and their antecedents. Decision Support Systems 44, 544–564 (2008)
15. Kline, R.B.: Principles and practice of structural equation modeling, 2nd edn. Guilford, New York (2011)
16. Kuan, H.H., Bock, G.W.: Trust transference in brick and click retailers: An investigation of the before-online-visit phase. Information & Management 44, 175–187 (2007)
17. Lewicki, R.J., Bunker, B.B.: Developing and maintaining trust in work relationships. In: Kramer, R.M., Tyler, T.R. (eds.) Trust in Organizations: Frontiers of Theory and Research, pp. 114–139. Sage Publications, Inc., Thousand Oaks (1996)
18. Pavlou, P.: Consumer acceptance of electronic commerce: Integrating trust and risk with the Technology Acceptance Model. International Journal of Electronic Commerce 17(3), 101–134 (2003)

19. Ranganathan, C., Goode, V., Ramaprasad, A.: Managing the transition to bricks and clicks. Communications of the ACM 46(12), 308–316 (2003)
20. Schreiber, J.B., Stage, F.K., King, J., Nora, A., Barlow, E.A.: Reporting structural equation modelling and confirmatory factor analysis results: A review. The Journal of Educational Research 99, 323–337 (2006)
21. Steiger, J.H.: Understanding the limitations of global fit assessment in structural equation modeling. Personality and Individual Differences 42, 893–898 (2007)
22. Tsai, H.T., Huang, H.C.: Determinants of e-repurchase intentions: An integrative model of quadruple retention drivers. Information & Management 44, 231–239 (2007)
23. Van der Heijden, H., Verhagen, T., Creemers, M.: Understanding online purchase intentions: Contributions from technology and trust perspectives. European Journal of Information Systems 12, 41–48 (2003)
24. Verhagen, T., van Dolen, W.: Online purchase intentions: A multi-channel store image perspective. Information & Management 46, 77–82 (2009)
25. Wen, C., Prybutok, V.R., Xu, C.: An integrated model for customer repurchase intention. Journal of Computer Information Systems 52(1), 14–23 (2010)

Interaction between Enterprise Resource Planning Systems and Organizational Culture

Carlotta Herberhold

University of Wuerzburg, Chair of Business Information Systems,
Wuerzburg, Germany
cherberhold@wiinf.uni-wuerzburg.de

Abstract. Building up on findings of IT culture research and ERP research, the assignment is to investigate organizational cultural variables impacting on changes through enterprise resource planning (ERP) systems. The study suggests seven cultural factors that are influence by ERP usage. The results are grounded in both a review of the literature on ERP, organizational culture, and an analysis of qualitative data collected from ERP experts. The findings will provide explanations that cultural factors should be considered when using an ERP system.

Keywords: Computer-augmented environment, organizational culture, Enterprise Resource Planning (ERP), user behavior, organizational behavior.

1 Cultural Battle with Enterprise Resource Planning

Enterprise Resource Planning (ERP) packages are implemented in enterprises with different organizational and national cultures. Culture refers to human practices, in contrast to conditions that are dictated by nature. Human values and behavior or communication patterns result from a cultural character. The latter decides whether certain behavior is accepted in a country or not. Thus, different cultures lead to diverse organizational structures in firms. They influence the culture of debate, the hierarchical understanding, organizational behavior, structures and information processing [3, p. 192; 4, p. 371].

A particular challenge is the introduction of new information systems and their changes since the users are not ready to give up old habits. Therefore, in order to maintain the old workflows they do not rely on the standard components. Ideally, however, the existing processes are not used as a basis for new target processes. Instead, the structure of the organization should base on the established processes of the standard software (reference processes) [12, p. 147,148; 3, p. 185,186].

The technical implementation of ERP systems is accompanied by a behavioral change in work processes, for example, that orders are processed accurately, material is dispatched on time or products are produced as needed [6, p. 231, 12, pp. 45, 55]. For management, it is important to realize that any change in the company also brings a change in corporate culture and sometimes requires it. Since the introduction of an

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 381–387, 2013.

ERP system has some significant impact on the way people work, the organizational culture is affected. Without a deep understanding of its impact on organizational culture, a company is unlikely to be able to develop with regard to the ERP system.

The maxim of each introduction or adaptation of an ERP system should not be that any resistance by the user is suppressed, punished and therefore useless—in other words a cultural battle—but that the consistent use of ERP systems improves processes and thus leads to a new company and organizational culture in the interest of the company and also its employees.

2 Research Gap and Research Questions

The objective of this study is to reveal the clash of cultural forces between the (cultural) practices embedded in ERP software and organizational culture and structure. The aim is to identify organizational culture values related to the meaning of practices embedded in ERP software (information, integration etc.). In detail, what characteristics, behaviors and attitudes counteract in regard to different organizational cultures with the features of an ERP system or what organizational cultures work easier with an ERP system.

Scholars looked at following aspects: Information system development, inter-organizational relationships, IT adoption and diffusion, IT management and strategy, IT use and outcomes, IT culture, IT impact on culture, organizational impact of IT. LEIDNER observes in her literature review that "a strong interest in understanding the relationship of IT and culture and determining how social groups interact with and apply IT in organizational settings" exists [4, p. 371]. Although culture has been cited widely as a challenge in ERP initiatives, and although many studies have considered the implications of organizational culture on ERP, no study addresses the change of organizational culture through an ERP system in a concrete and observable manner (see figure 1). Studies on ERP and culture generally include the adoption of ERP in different countries in order to explain the role of culture in implementing ERP systems. This article adds to the literature by identifying organizational culture variables that change through ERP systems in an organization.

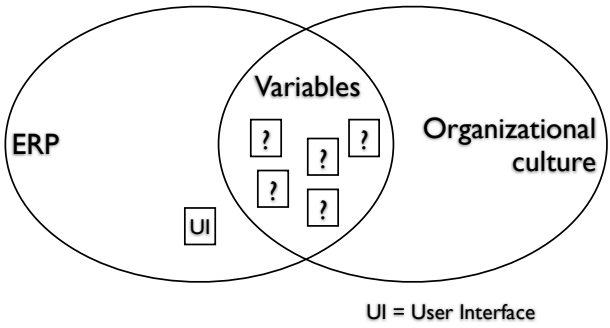


Fig. 1. Interaction between ERP and Organizational Culture

Specifically, the following research questions (RQ) will be answered: What organizational culture variables change by ERP?

This aspect will contribute to the understanding of ERP systems and the associated organizational changes in companies.

3 Research Design

The author pursues to contribute to IT culture research by identifying organizational culture variables that change through ERP systems in organizational settings. An explicit study of variables in order to identify requirements for organizational culture through ERP has not been made yet in any work. For this reason, it is first necessary to develop an understanding of this topic. A qualitative approach, as opposed to a quantitative method, is a suitable research approach, if a topic has not been understood sufficiently. They offer the variety and wealth of data to get an understanding of the context to be examined [5, pp. 19-22].

In the first part, the author strives to provide components and requirements of the ERP system that is necessary for the analysis. The important point is that only components related to the organization are looked at.

The second part has the objective to identify organizational culture variables that change through ERP systems in an organization. In order to understand the interaction between culture values and ERP, this paper reports on an exploratory interview study with ERP experts that identify these values [13, pp. 76; 9] (see table 1). The interview guide provides questions about organizational structure, processes and data as well as information usage and usage in general. The survey of these people allows to get a better understanding of the ERP-cultural relationship and thus to detect causal relationships than it would be possible through a survey. The selected group of experts has years of experience and knowledge in dealing with the implementation and use of ERP software in companies [13, 14].

The third part has the objective to measure the variables before and after ERP introduction. The case study approach is chosen as research method. The case study approach enables to explore and point out those questions with the concurrent lack of knowledge about identify organizational culture variables that change through ERP systems (see table 1). The case studies are based on observing different units of analysis in its natural setting. The filtered data of numerous single observations (each building one case study) is then combined and examined against the study objectives. Propositions are deduced from the results and can, at a later stage, be verified by a large-scale quantitative survey.

The overall objective of the empirical part is to identify requirements for organizational culture through ERP systems based on organizational culture theories (e.g. Schein) and provide recommendations to support the ERP change management process through adaptation of organizational culture (ERP culture alignment).

Table 1. Research process

Empirical study	Objective(s)	Method
I (present paper)	Provide ERP components and requirements (unit of analysis)	Literature review
II (present paper)	Identify organizational culture variables that change through ERP systems in an organization (RQ 1).	Interview study (ERP experts)
III	Measurement of the variables before and after ERP introduction (e.g. MM-Module, OTC process, PTP process).	Embedded case study (different units of analysis in an organization and different experts)
IV	Identify requirements for organizational culture through ERP based on organizational culture theories (e.g. Schein) (RQ 2).	Argumentative-deductive
V	Recommendations to support ERP change management process through adaptation of organizational culture (ERP culture alignment)	Argumentative-deductive

4 Literature Review

4.1 Enterprise Resource Planning

Business software or enterprise systems are different names for Enterprise Resource Planning (ERP) systems. An ERP system is a special type of information systems, which are defined as socio-technical-systems. It is used for the quantitative and time-oriented scheduling of equipment, personnel and capital. The system controls the process depending on the material types, document types, groups of accounts or payment but independent of the people who work in the various areas of the company. Compared to other information systems, ERP is characterized by best-of-breed business processes. These standard software systems were designed and developed for the interaction between different task fields from the very beginning [12, pp. 147, 148].

For the whole research process, the following four components and requirements of ERP system are relevant [3,11, 12]:

- **(Organizational-) structure.** When using an ERP system it is necessary to illustrate the organizational structure of the enterprise in the structure of the standard business software. That means it is necessary to transfer the real organization into system elements like e.g. client, company code, plant and storage location. Besides this the company is divided into functional areas, like materials management, production planning, sales and distribution, financial accounting, controlling etc.

Also, in order to capture process information and the logical structuring of the workflow various functional areas within the company have documents as control and steering elements.

- **Processes.** The usage of ERP systems is meant to improve business processes. These improvements are related to e.g. activities, division of labor, responsibilities and deliverables. Besides, ERP means automation of functionality. For example, transformation rules calculate and derive something in the pricing process.
- **Common data.** ERP systems also share a common database, which enables consistent organizational data collection and processing. A standardized way of communication and information distribution is meant by ERP.
- **User Interface (UI).** The UI is the interface between system and user. Thus, a user interface for humans must be adapted to their needs and abilities in order to be useful to them. Usability criteria analogous to websites are colors, grouping of content and spatial distribution. The culture takes an affect on how the user is dealing with the product and why people from different cultures have different ideas of usability.

Since this paper is about dealing with ERP in the context of different organizational culture dimensions, this rather technical aspect of the UI is not further included in the analysis. It also rather concerns national-specific differences than organizational culture. The organizational culture influences how users work with the ERP system and why people have different understandings of using the system.

4.2 Organizational Culture

Companies differ not only by location, size, products and their success; they are also characterized by different objectives, principles, manners and behaviors. Not only the country in which they reside shapes corporate cultures. In turn, corporate culture is the basis for rules, procedures, working conditions, habits, decision-making, knowledge sharing, and information usage [2].

In his organizational culture framework Edgar Schein (1988) describes culture as follows: „Culture is a pattern of basic assumptions, invented, discovered, or developed by a given group, as it learns to cope with its problems of external adaptation and internal integration, that has worked well enough to be considered valid and, therefore is to be taught to new members as the correct way to perceive, think, and feel in relation to those problems." His model illustrates culture on three different levels, which influences each other: Visible organizational structures and processes, values and beliefs [10, p.9].

Other frameworks include dimensions or values to detect organizational culture. With information technology, in particular the following content areas of corporate culture, denominated as cultural dimensions are put in contact: target orientation, result orientation, cost orientation, open communication, process orientation, ability to learn and adapt. These dimensions are partially interrelated or interdependent or mutually dependent [1, 4, 8, 9].

5 Preliminary Results

The interview study was organized on the basis of the factors identified in the literature review. The review leverages these components of ERP: structure, processes and data. The results of the literature review were then discussed in semi-structured interviews with ERP experts who have experienced a high number of ERP implementations and change processes. The interview results revealed the following organizational culture variables that change through ERP systems:

- **Decisions.** In order to make a decision managers need information about all the different enterprise divisions. Through ERP this process has changed. Also the users are confronted with a new situation. For instance the system has automated some process steps changing the work environment of the user. This may influence the way of decision making.
- **Transparence.** An ERP system establishes the principle that information is centrally available. It enables the transfer of information that was previously available only in one place. However, this requires open communication and the willingness to share information. The interview results showed that on the one hand users were sometimes exchanging more information among each other and on the other hand higher transparence were frightening them in the beginning.
- **New skills of user.** With the introduction of an ERP system the user has to learn how to use the system technically, but he also needs to learn new things like thinking in a structured and logical way, accuracy, unambiguity and the ability to abstract.
- **Evaluations.** ERP enables to evaluate transactional data i.e. business transactions. A customer-oriented, objective and logically correct process control takes center stage.
- **Responsibilities and accountabilities.** Working with ERP systems is goal-oriented, structured and systematic according to schedule. Organizational units and employees are assigned clear roles and responsibilities. An organizational culture that includes this already works—a positive effect on the ERP application.
- **Process Owner with a clear mandate and responsibility.** ERP systems provide the basis to overcome divisional thinking and going over to enterprise-wide process thinking. The central information management and consultative information processing enables process integration and collaboration among departments with cross-functional teams. This will only work with a process owner provided with power of decision.
- **Avoidance of central data storage.** ERP systems are based on the idea of data integration. Important information is collected, stored and processed at the origin. This is important for the smooth running, but also for a high level of information security (compliance). For example, every movement of goods triggers changes in accounting and cost accounting. The scheduling and procurement costs, inventory levels and the cost of capital are controlled. Causalities become clear. This new way of data and information usage is new to the user.

Interaction between ERP and Organizational Culture. The overall goal of this research was to show how organizational culture values are reflected in relation to the ERP systems. The challenge to use a business information system in the correct way is not only a technical matter but, like the interview results show, a matter of interaction of system requirements and organizational culture like decision making, knowledge sharing, information usage, structures and processes, values as well as beliefs (variables).

Further research should answer the question, how organizational culture influences the understanding for changes induced by ERP. The objective is to design this interaction into practical innovation for a better adoption of ERP systems. Also a organizational culture should be described which fits best the intended usage of ERP systems and how an iterative IT-based change leads to a new ERP culture with its own information and IT values.

References

1. Hofstede, G., et al.: Measuring Organizational Cultures: A Qualitative and Quantitative Study Across Twenty Cases. Administrative Science Quarterly 35(2), 286–316 (1990)
2. Kobi, J.-M., Wüthrich, H.A.: Unternehmenskultur verstehen, erfassen und gestalten. Verlag Moderne Industrie, Landsberg/Lech (1986)
3. Krumbholz, M.: The Implementation of Enterprise Resource Planning Packages in Different Organisational and National Cultures. Information Systems (26), 185–204 (2001)
4. Leidner, D., Kayworth, T.: A Review of Culture In Informations Systems – Towards a Theory of Information Technology Culture Conflict. MIS Quarterly (30), 357–399 (2006)
5. Mayring, P.: Qualitative Inhaltsanalyse: Grundlagen und Techniken. Beltz Verlag. Weinheim (2010)
6. Merkens, H.: Interkulturelles Management: Die Herausforderung des 21. Jahrhunderts. In: Interkulturell denken und handeln: Theoretische Grundlagen und Orientierung für die gesellschaftliche Praxis, pp. S.229–S.239 (2006)
7. Richter, T., Adelsberger, H.: Kulturspezifische Untersuchungen in der gestaltungsorientierten Wirtschaftsinformatik: Bestandsaufnahme und Analyse des derzeitigen Umgangs mit "Kultur". Universität Duisburg-Essen (2011), http://duepublico.uni-duisburg-essen.de/servlets/DerivateServlet/Derivate-28380/017_FinalDuePublico_WI_Umgang_Kultur_Richter_Adelsberger.pdf
8. Sackmann, S.A.: Welche kulturellen Faktoren beenflussen den Unternehmenserfolg? Universität der Bundeswehr München, Neubiberg (2006)
9. Sagiv, L., Schwatz, S.H.: Cultural values in organisations: insight for Europe. European J. International Management 1(3), 176–190 (2007)
10. Schein, E.H.: Organizational Culture and Leadership. John Wiley & Sons, New York (2004)
11. Thome, R., Hufgard, A.: Continuous System Engineering. Oxygon Publisher, Munich (2006)
12. Thome, R.: Grundzüge der Wirtschaftsinformatik – Integration der Informationsverarbeitung in die Organisation von Unternhemen. Pearson Studium, Munich (2006)
13. Walsham, G.: Interpretive case studies in IS research: nature and method. European Journal of Information Systems 4, 75–81 (1995)
14. Yin, R.K.: Case Study Research: Design and Methods. SAGE Publications, Thousand Oaks (2003)

Reconfiguring the Corporate and Commons: Mobile and Online Activism as a Form of Socio-technical Design

Constance Kampf

Business & Social Sciences, Aarhus University, Aarhus, Denmark
cka@asb.dk

Abstract. This paper describes a model for understanding how digital activists contribute to (re)defining the ways in which socio-technical interactions are (re)designed at the Macro-level. This macro level is explained in three parts—1) the cultural production of a new form of practice, 2) an active, contested knowledge process, and 3) as producing knowledge about the interface between society and business. Two kinds of digital activists are examined—critical digital activists and activists from within the system.

Keywords: Socio-Technical Design, Culture, Knowledge.

1 Introduction

What is digital activism and why does it matter? Since the Lotus Marketplace and Clipper Chip online protests [1], technology has begun to shift the balance of power towards consumer activists who are able to use it effectively to interrupt product development and suggest changes for corporate strategy. With critically focused activists such as Wikileaks, the hacker's group Anonymous, and digital artist activists such as Ubermorgen [2], technology is being used to engage with power structures and shift the balance of power between the corporate and the commons.

Both critical Digital activism and "activism from within the system" can be seen as constituting a force for (re)design of social technical interactions. [3] calls for two levels in our understanding of socio-technical design, a micro level focused on user level interaction, and a macro-level focused on society level interactions. These society level interactions between organized groups of activists and corporations work to shift the attention of business from itself and its own goals to surrounding issues.

This paper focuses on discussing these two approaches to Internet activism as emerging forces working to shape and shift the balance of power and the face of technology based interactions between corporations and the commons as a form of socio-technical design. Under-standing these two approaches to Internet activism helps demonstrate ways in which activists towards business engage in socio-technical (re)design at the macro level.

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 388–395, 2013.

2 Method

This paper uses argument and example to lay out a theoretical frame-work for understanding how different forms of online activism provide forces shaping socio-technical (re)design at the macro level. Both primary and secondary data are used to point out patterns among cases of Internet activism, and reflect on their implications for explaining how different kinds of activism function as a force shaping the (re)design of socio-technical interactions. Primary data consists of social media postings in Blogs, YouTube, Facebook as well as websites collected from 2008-present. Primary data is focused on an NGO—CarrotMob, and a B corporation, the Good Guide.

This primary and secondary data is approached through a multidisciplinary lens combing sociotechnical design, organizational theory and sociological concepts. These concepts draw from two distinct levels of understanding the interaction between people and technology. First, from a workplace level, applying socio-technical design and organizational theory we can find a perspective on knowledge that is both socio-technical and process oriented in nature. Adding a more general level, that of society, we can expand that perspective to include an epistemic or knowledge producing space that affects and interacts with culture at the societal level.

This lens draws on the following concepts. From a traditional workplace focused perspective on socio-technical design, Suchman et al [4] describe systems development as cultural production of new forms of practice. From the field of Management, Blacker argues for under-standing knowledge as an active process in a culturally located system [5]. According to Blackler, this process is contested, provisional, and pragmatic. In the case of a Macro approach to Socio-Technical design, this culturally located system can be understood as an epistemic, or knowledge producing. Knorr Cetina puts forward the notion of a general knowledge culture in which specific knowledge processes are embedded [6], from the field of sociology. She uses the term "epistemic" or knowledge-producing to describe a knowledge culture, focusing on the ways in which knowledge is produced and reproduced through what she terms "epistemic circuits" connecting actors and technologies. We can link these three perspectives together for understanding a macro level of socio-technical design. This macro level can be constructed through examining Blackler's active, contested knowledge processes as part of a global epistemic culture following Knorr Centina's sociological frame of reference, in which activists are engaging with social and mobile media to produce contested, provisional and pragmatic knowledge about the interface between society and business. This use of social and mobile media can be informed by Suchman et al's connection between technology and cultural production to a global, macro-level for socio-technical design [3].

To operationalize this multidisciplinary lens, Critical digital activism and "Activism from within the system" will be examined through 1) the cultural production of a new form of practice [4]; 2) as an active con-tested knowledge process [5]; and 3) as producing knowledge about the interface between society and business [3,6]. These 3 aspects work together to demonstrate how each type of activism can be understood as a force for (re)design of socio-technical interactions through linking together practice, process and the larger picture of the interface between society and business.

3 Critical Digital Activism as a Force for (re)design of Socio-technical Interactions

Critical digital activism reconfigures the commons and the corporate in unexpected ways. The technologies of email, websites, and social and mobile media open up spaces for critics of business strategy to be heard and gather followers with global reach.

3.1 The Cultural Production of Reach in the Public Critique of Business Strategy

In the late 1980s, the movement of email, bulletin boards, and discussion lists from the workplace to society enabled a shift in the ability for large groups of protesters to communicate with each other and companies. An important type of protester was the consumer activist. Lotus Marketplace is an example of a product that was widely protested through these new channels [1]. In the end, consumer actions led to change in company strategy and the product was cancelled before it reached the market.

Another type of critical digital activist is the artist-activist. The artist-activists do not always begin as activists, but can react to the pressures of companies when they feel the company is inappropriate. An example of this can be seen in the etoy war, arguable the most expensive art performance in history. The digital artist activist group etoy.com was legally pressured by a U.S. new technology startup, eToys.com to stop using their domain name because it was too similar to etoys.com and etoys.com customers who mistyped the URL were offended by etoy.com content. First the company sent them a request for their domain name, and when it was refused, the company eToys.com went to court and won an injunction preventing the digital artist group etoy from using their domain. The group was forced by US legal action to comply, and they felt it was unjust because they had owned the domain etoy and built their brand in Europe before the company eToys.com was started. In reaction, they opened a new domain, Toywar.com, and started "involving the enemy in an insane ToyNam situation." [2]. The new domain, Toywar.com, supported a game that awarded points for actions that lowered Etoys.com stock. The artists recruited 1799 hackers to play the game. The result of this art performance was bankruptcy for eToys.com, which was sold at a loss of billions of dollars [2]. Here, the artists not only affected the business strategy, but they felt attacked by it, and so they coordinated a responding attack. The reach of these artists enabled them to affect eToys.com in an unprecedented manner.

3.2 Critical Digital Activism as an Active, Contested Knowledge Process

In the cases of Lotus Marketplace and the eToy war, the activists engaged in communicating ideas which addressed a contested boundary between business and society. In the case of Lotus Marketplace, the product itself was selling personal information that was already available publicly, but packaged in a CD-ROM format that (for that point in time) would make it easier for marketers to target consumers. In the late 1980s,

consumers were upset about a product that they understood would affect their privacy. Thus, protesting Lotus Marketplace as a product could be understood having roots in an active discussion. Gurak describes the discussion as taking place over a period of several months, during which computer specialists and consumers engaged in technical arguments about the product. She describes postings as having "a range of functions: to inform, to debate the protest, to debate the product, to mobilize, to divulge "secret information" and to discuss other privacy related issues." Thus, the activism had roots in a contested discussion about the nature of privacy at the interface of business and society [1, p.27].

On the other hand, the digital activists focused on issues of ownership and legal rights, contesting the perspective of the corporation and the US courts by statements on their website telling the story, as well as by actions of creating a space where the corporation was not prioritized because of its' resources, but rather reconfigured to become the enemy of free art and free speech. Here, the contested understandings of the situation were not discussed as much as enacted in the art performance, described as a virtual "ToyNam" by the artists [2]. Because the force of courts had set up the knowledge that a company has the right to take away a domain that had been held by another group before the company was created, the artists chose to act out their disagreement rather than discuss it, and within the framework of the game offered different options for supporting "toy soldier" hackers to lower the stock of Etoy.

Both of these cases demonstrate that activism occurs within an active and contested space, and the duration of the Lotus Marketplace discussions over a period of several months, as well as the duration of the ToyNam performance over about a year indicate that the activism was not a single effort, but rather an ongoing, engaged process.

3.3 Critical Digital Activism as Producing Knowledge about the Interface between Society and Business

Both Lotus Marketplace and the ToyNam art performance by the etoy artist collective offer examples of activists focusing on the interface of business and society as it is shaped and (re)shaped by technology. Each case focuses on a different key issue. For the Lotus Marketplace product, privacy and consumer perception of privacy became a key issue. At the time, the knowledge produced around privacy was communicated through then current technology: usenet groups, bulletin boards, MOOs and MUDs, and email. This knowledge addressed but did not resolve issues related to privacy, and we are still dealing with them in different forms today. For the ToyNam art performance, the issue was virtual domain ownership—do companies have the right to interfere with previously owned domains because it can affect their business? The response of the artist activists set a clearer boundary by creating the knowledge that a group of unhappy digital artists and their supporters had the ability to affect the success of a large and well funded business which had, in effect, attacked the digital artists' brand.

Both of these issues—issues of privacy and domain ownership, can be seen as cultural in nature. The ways in which people choose to question or accept privacy standards become cultural norms over time. The question is to what extend are we

drawn into adapting our cultural practices to fit the needs of the business? For example, recent Google and Facebook practices have taken the notion of transparency and turned it over towards the consumer, setting up consumer transparency as a key issue rather than consumer privacy.

4 "Activism from within the System" as a Force for (re)design of Socio-technical Interactions

"Activism from within the system" reconfigures the commons and the corporate around a principal focus or single topic approach, and is designed to push an agenda in way that creates a "win-win" situation for both business and society. These digital "activists from within the system" use technology to facilitate changes within the current corporate-commons balance. We will examine two examples: CarrotMob, an NGO led by Brent Schulkin, and Good Guide, a B-Corporation led by Dara O'Rourke. These organizations use technology and social media to build awareness in the commons around the issues of Corporate Social Responsibility (CSR) and sustainability, and put pressure on corporations to change by engaging them in win-win scenarios.

4.1 The Cultural Production of Corporate Social Responsibility in the Public Critique of Business Strategy

Like the critical activists, activists from within are also engaged in using or appropriating technology for cultural production. In this case though, they go beyond the cultural production of reach already established by earlier digital activist movements to the cultural production of CSR. The framework for CSR is contested, and corporations work at sensegiving [7] through their communications as much as NGOs, and government bodies also work to define it. In producing cultural norms for CSR, CarrotMob and Goodguide each focus on a different definition. For CarrotMob, the cultural production of CSR is done using video campaigns in YouTube to create a dialog with local business owners and activists around sustainability issues. The key video of the organization is a cartoon that shows business as polluting the environment, and not responding to government, but rather to people who have pooled their resources enough to gain corporate interests [8]. The goal of the organization is to create a new cultural norm or resource through the activity of organizing purchases and negotiating with businesses to engage in CSR such as reducing their environmental footprint. These issues engage with different cultural norms in the different locations of the CarrotMob events. For example, the CarrotMob event in Jönköping, Sweden, focuses on purchasing fair trade goods in local grocery stores [9]. In contrast, the CarrotMob event in Eugene, Oregon has a number of videos that focus on reducing waste, energy use, and engaging a local supply chain in local coffee shops [10]. There are currently 328 videos from around the world that show aspects of CarrotMob events. Each of these videos is localized to create a cultural production of the practice of coordinated purchasing appropriate to the local cultural context.

The Good Guide also engages in a different kind of cultural production of CSR. It focuses on the coordination and interpretation of scientific data about products to rate them with respect to three categories of CSR—health, environment and community. The Good Guide is unique in its' focus on health as a component of CSR for products. To create this type of activism from within the system, the Good Guide is a for profit company with a NGO style mission. Thus, companies hire Good Guide to feature their products in Good Guide's information system. This system is available as a database from their website, as well as an iPhone app which scans in product bar codes at the point of purchase. The Good Guide app then shows their rating system for each of the their three key categories, allowing the user to find more information about the product's ingredients, as well as keep track of how socially responsible their purchases are. The rating system is a 10-point scale, with 4 shades of color to remind consumers about their choices. These shades are green for products that are good, and shades of brown, orange and red for products that have scientific data that demonstrate issues in one of the three areas. Thus, by using the Good Guide smart phone app or database to evaluate products before purchase, consumers enact responsible purchasing, which then ostensibly puts pressure on companies to produce more responsible products.

4.2 "Activism from within the System" as an Active, Contested Knowledge Process

Both CarrotMob and GoodGuide offer different approaches to opening up conversations around CSR. For CarrotMob, these different approaches to CSR knowledge can be seen in the video content, with different sponsors emphasizing various aspects of CSR from fair trade products, to waste reduction and local supply chains, to workers rights. Although the CarrotMob web presence does not specifically offer a space for discussing issues related to sustainability, the format for Carrotmob events is open enough to allow organizers to focus in on issues, which they find important and relevant in their context. Thus, the collection of CarrotMob videos operates as a contested knowledge process in which consumers from around the globe are engaging in building definitions of CSR appropriate to their local contexts. For GoodGuide, these approaches focus on coordinating and organizing scientific research in their database, as well as encouraging consumers to interact with government regulators with issues related to health. In the website database, consumers are offered the ability to recommend, suggest avoiding, and/or comment on their experience with each product. This opens up a formal space for dialogue and allows for an active, contested knowledge process to build knowledge about each product while including multiple perspectives.

4.3 "Activism from within the System" as Producing Knowledge about the Interface between Society and Business

CarrotMob and GoodGuide both work to produce knowledge about how consumer action and dialog affects business strategy in win-win activist scenarios. In addition, they work to produce knowledge about ways in which consumers can engage in

consuming with responsibility in order to encourage companies to offer opportunities for responsible/sustainable consumption. Thus, the interface between society and business in these cases focuses on sharing knowledge about products, and brings a critical perspective to consumption. The idea behind this type of activism is then, that this communication loop between consumers explaining their strategy for sustainability, or their needs for a product that protects their health as well as the community and environment and corporations will lead to (re)definition of CSR as well as a (re)definition of the roles for consumers and business to engage in co-producing CSR through their actions, which in turn affects the possibilities for knowing how consumers and businesses should interact at the interface between society and business.

5 Conclusion

In conclusion, this model for understanding the effects of digital activism on culture can be seen through the overall idea of a Macro-level for socio-technical design. This Macro-level focuses on cultural changes in society. It can be operationalized through the three aspects used to examine digital activists—1) cultural production, 2) an active, contested knowledge process, and 3) producing knowledge about the interface between society and business. These three aspects of a Macro-level for socio-technical design were examined using examples from two kinds of digital activists—critical activists and "activists from within". The activists in both categories examined here followed similar processes, which can be understood as a macro level for socio-technical design. Each group of activists worked towards engaging in cultural production, an active contested knowledge process and producing knowledge about the interface between society and business. By following the Macro-level for socio-technical design, one can see a way of explaining how technology affects our possibilities for action, activism, and engaging in both processes of creating knowledge and producing culture. In order to develop a more nuanced understanding of a macro-level for socio-technical design, connecting these three aspects to definitions of culture in global contexts would be a useful next step.

References

1. Gurak, L.: Persuasion and Privacy in Cyberspace: The Online Protests over Lotus Market-Place and the Clipper Chip. Yale University Press (1997)
2. Kampf, C., Cox, G.: Using Digital Art to make the tensions between Capital and Commons Transparent: Innovation in shaping knowledge of Internet business practices as a form of cultural knowledge. In: Cultural Attitudes Towards Communication and Technology Proceedings, Aarhus 2012. Murdoch University Press (2012), http://blogs.ubc.ca/catac/proceedings/proceedings-2012/
3. Kampf, C.: Revealing the Socio-technical Design of Global e-Businesses: a case of digital artists engaging in radical transparency. International Journal of Sociotechnology and Knowledge Development 4(4) (2012)
4. Suchman, L., Bloomberg, J., Orr, J.E., Trigg, R.: Reconstructing Technologies as Social Practice. American Behavioral Scientist 43, 392 (1999)

5. Knorr Cetina, K.: Culture in global knowledge societies: knowledge cultures and epistemic cultures. Interdisciplinary Science Reviews 32(4) (2007)
6. Morsing, M., Schultz, L.: Corporate social responsibility communication: stakeholder information, response and involvement strategies. Business Ethics, A European Review 15(4), 323–328 (2006)
7. Gioia, D.A., Chittipeddi, K.: Sensemaking and Sensegiving in Strategic Change Initiation. Strat. Mgmt. J. 12, 433–448 (1991)
8. Shulkin, B.: How organized purchasing can change business. YouTube Video (2008), http://www.youtube.com/watch?v=AkGMVQl8_jI (accessed March 1, 2013)
9. Rättvis handel Jönköping. Carrotmob premiers in Jönköping (2010), http://www.youtube.com/watch?v=pZ_Xlrn2PQU (accessed March 1, 2013)
10. Eugene Carrot Mob (2990). Eugene CarrotMob Teaser, http://www.youtube.com/watch?v=pZ_Xlrn2PQU (accessed March 1, 2013)

Social Media in the Product Development Process of the Automotive Industry: A New Approach

Andreas Klein and Götz Spiegel

Spiegel Institut Mannheim GmbH & Co. KG, Schwetzinger Str. 32, D-68782 Brühl, Germany
a.klein@spiegel-institut.de

Abstract. This paper introduces a new methodology for implementing social media monitoring into an important stage of the innovation process within the automotive industry — the prototype stage. The information gathered on social media channels was used for project management, quality management and idea-generation purposes in order to support the parallel development of future generations of electric vehicles. As the project is still ongoing the first general results which show the effectiveness of the methodology can be presented in this paper.

Keywords: Social media monitoring, electric mobility, electric vehicles, product development, social networking sites.

1 Introduction and Theoretical Background

In the project this paper describes social media sources are used in an automotive context in order to support a real-world prototype stage. As the project and the monitoring process are still ongoing and final results will probably be available in autumn 2013, the focus of this paper is on the innovative methodology and a few preliminary results.

The importance of social media as a communication channel for companies is obvious, especially after the spectacular IPO of Facebook at the beginning of 2012. But social media is not only about social networking sites as Kaplan et al. [1] already pointed out. It also includes blogs, virtual worlds, games as well as collaborative projects like Wikipedia and most importantly content communities. Commonly acknowledged by the scientific literature is the usage of social media tools for marketing purposes, e.g., the introduction of the new BMW X1 [2]. The emphasis of such activities lies more or less on an expected increase of public attention and on targeting new customer groups. But often it is also used as an instrument to gain information about the requirements and wishes of the customers. This scientific analysis of internet communities in order to learn more about the members, their mode of communication and the subject of their discussion was very much influenced by the work of Robert V. Kozinets who used ethnographic methods for online communities [3], [4]. These methods are widely applied, e.g., to learn about brand communities of the automotive industry [5]. Social media sources seem to be particularly suited to give input for the innovation process — from the idea creation to virtual product tests [6].

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 396–401, 2013.

2 Project Description

In 2007 BMW Group´s task force "project i" decided to include electric vehicles at a larger scale into the product portfolio [7]. According to the communicated timeline, BMW is expecting its first series-production full electric vehicle with the BMW i3 at the end of 2013. To achieve this goal, electric prototypes on the basis of existing cars with internal combustion engine were produced and tested in order to gain more experience with electric vehicles. The first converted concept was the MINI E. Several hundred units were produced and internationally tested by private and fleet users between 2009 and 2012 accompanied by an all-embracing scientific research approach in order to learn about the acceptance, requirements and barriers for electric mobility [8]. In this context the BMW Group was closely supported by the German research institute "Spiegel Institut Mannheim".

The second converted vehicle concept introduced by the BMW Group is the BMW ActiveE, derived from the 1-series. In this vehicle, series components for the following series-production electric vehicles like the BMW i3 have already been integrated. After the usual intensive internal testing by BMW regarding technological and safety issues, the car was ready for being shipped to the customers. About 700 units have been shipped to the United States since January 2012 and handed over to customers for 24 months. These pay a monthly leasing fee of 499$ [9].

3 Research Methodology

Because of the nature of the BMW ActiveE trial — the test of a near-to-market technology while the mass-market product is at its final stages of development — two aspects are of major importance for scientific research:

- quality management aspects in order to identify potential product, service or usability problems
- idea generation from the driver community about potential product improvements for enhancing the user experience.

Therefore the research methodology has to meet some requirements. It needs to be flexible and any critical findings need to be communicated to the development departments instantly. Furthermore, it has to be considered that the time of occurrence of quality management aspects cannot be predicted as an electric vehicle is dependent on outer conditions like charging infrastructure, climate conditions etc. Additionally even single problems might have a big impact on the mass-market product and need to be reported to the development teams.

Considering these requirements, in addition to a "classic" market research approach, an additional social media research strategy was developed for the BMW ActiveE. A primary research revealed four adequate social media sources for the later analysis (see table 1). Besides the official BMW ActiveE web forum three additional discussion groups have been founded by BMW ActiveE drivers on social networking

sites. Most members either drive a BMW ActiveE or are very much interested in the product. Many drivers are members of the forum as well of the discussion groups at the same time. All discussion groups are open and can be fully read by the public.

Table 1. Sources for Social Media Analysis[1]

Name	Members	Number of posts	Founder
The BMW ActiveE Forum	1.806	9.709	BMW Group
BMW ActiveE	384	24.367	Private Users
BMW ActiveE San Diego	17	91	Private Users
BMW i3 – ActiveE – Megacity EV	75	139	Private Users

The core instrument is a data base with all posts and comments assigned to 55 different categories. On a weekly basis all new posts are transferred manually into the data base. The categories include a simple sentiment analysis (positive/neutral/negative) and a more detailed view on all vehicle related aspects. A special focus lies on new aspects that bring electric vehicles to the customers, e.g., regenerative braking or battery

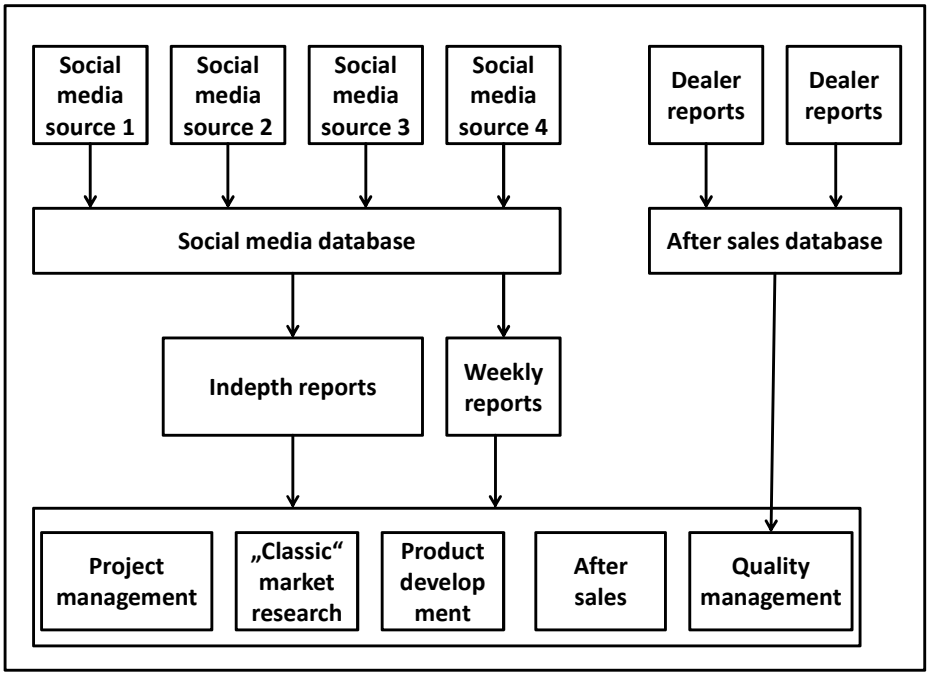


Fig. 1. Social Media in the research process

[1] In February 2013.

lifetime. The database provides the possibility to run in-depth data analysis for certain topics and to identify trends. In order to track the project progress, to early identify critical project developments, and to generally get direct insight into the topics of current users' discussions, a weekly report including a summary of the most important discussions and reported problems on the monitored social media channels is generated.

While the most important stakeholders for the results within BMW are the development and quality management departments, the weekly reports also provide for the project management and after sales departments the possibility to gain nearly instantly direct insight into the customers' most important discussed issues and thus to react if necessary (see Fig.1.). And finally, the social media insights also provide a potentially valuable source of qualitative data for the development of the accompanying "classic" market research.

4 Results

While comparing the user activity on the official BMW ActiveE forum and the most important discussion group, it is obvious that although more people have subscribed to the forum the more active communication is on the social networking group.[2] The discussion on the four social media channels can roughly be divided into three categories:

- social communication about events, e.g., motor shows/meetings with other users
- discussion about problems/product-related questions/help from other drivers.
- statistics about driven miles/used electricity for BMW ActiveE.

While each of the categories applies to the forum and the social networking groups equally, the direction of the discussion on the social networking groups is more within the users, while the posts on the forum mostly aim at BMW, e.g., desired improvements for future vehicles. From the sentiment analysis it is difficult to draw a "balanced" picture of the advantages and disadvantages of the product as positive attributes are mostly expected while negative attributes contain a certain news value and are therefore reported more often. So a proper assessment of the vehicle features has to be carried out by the accompanying "classic" market research.

Nevertheless, it is a good tool to identify potential product problems. To give an idea of the effectiveness of the methodology for quality management, the findings from the social media sources have been compared to the information collected by the conventional quality management process in a BMW internal after sales database. The after sales database is based on the service data from the dealers. So if BMW ActiveE customers encounter a vehicle problem and ask the dealers for help, the dealers would use the after sales database for support.

Within a monitored period of the first nine months the BMW quality management team thereby identified 39 product-related problems which occurred at least twice. The social media analysis meanwhile revealed 35 different product-related problems. These problems did not all demand a stopover at the garage but influenced the

[2] Nardi et al. give a good description of the motivation for activity on social media sites [10].

product experience in a negative way. About 32 problems were already covered by the dealer reports. One reported problem with the vehicle door could not be validated by the BMW engineers. Two problems had only been identified with the support of the social media analysis and could be tackled by quality management measures. One serious problem was already known but the frequency of occurrence was a lot more than previously expected by the experts.

In order to gain new product ideas, an in-depth analysis of the four social media sources identified 23 new ideas or suggestions for the BMW ActiveE or electric vehicles in general. Thereof 12 ideas have already been part of the product strategy for the successor, the i3. About eight ideas are currently in a review process by the development departments. Only three ideas had been declined because of not fitting to the product strategy.

Finally, the social media monitoring also turned out to be a valuable source for the project management. So far the weekly report is currently subscribed by more than 130 involved experts from different departments throughout the entire BMW Group, involving project management, development, communications and sales.

5 Review of the Social Media Monitoring Research Methodology

To summarize, the described methodology has a number of advantages and disadvantages.

Advantages:

- nearly real-time feedback on the product from the users' perspective over the trial period
- non-intrusive methodology; users explain problems in their own language; topics are set by the users themselves
- cost-effective compared to other near real-time "classic" market research methodologies
- useful way to gather qualitative feedback as input for a later broader quantitative market research design.

Disadvantages:

- no exact quantitative results about the occurrence of problems or about the opinion on ideas possible, because only a subset of the users post on social media channels
- no "balanced" picture of the product; the communication is rather more problem-driven.

6 Conclusion

Although social media monitoring during the product prototype stage contributed a significant part to the collected information, it did not provide the full picture. It is

therefore rather an important element than a full replacement of a broader "classic" research strategy. Furthermore, it was very helpful to do the research with a high-involvement product such as a new electric vehicle which also attracts rather tech-savvy customers who are more likely to use social media channels. Finally, the limited number of participants in the trial also limited the amount of data, which facilitates the analysis process.

Acknowledgments. The authors would like to acknowledge the BMW Group for allowing us to listen to the customers' voice, especially Roman Vilimek and Andreas Keinath for their guidance and support throughout the international MINI E and ActiveE projects. And, of course, we would like to thank all the BMW Electronauts who supported the entire development process of the i3 and its successors with their valuable feedback. Finally many thanks also to the Spiegel team especially Alexander Stech for summarizing the weekly reports and designing the database, and Julia Messerschmidt for her valuable in-depth analyses.

References

1. Kaplan, A.M., Haenlein, M.: Users of the world, unite! The challenges and opportunities of Social Media. Business Horizons 53(1), 59–68 (2010)
2. Douglas, T., Lorenz, M., Oheimb, C.: Pre-Kommunikation im Internet — Der BMW X1 trifft Kunden im Social Web. Marketing Review St. Gallen 26(4), 11–15 (2009)
3. Kozinets, R.V.: E-tribalized marketing? The strategic implications of virtual communities of consumption. European Management Journal 17(3), 252–264 (1999)
4. Kozinets, R.V.: The Field Behind the Screen: Using Netnography for Marketing Research in Online Communities. Journal of Marketing Research 39(1), 61–72 (2002)
5. Füller, J., Matzler, K., Hoppe, M.: Brand Community Members as a Source of Innovation. Journal of Product Innovation Management 25(6), 608–619 (2008)
6. Füller, J., Bartl, M., Ernst, H., Mühlbacher, H.: Community based innovation: How to integrate members of virtual communities into new product development. Electronic Commerce Research 6(1), 57–73 (2006)
7. BMW Group: BMW i. Born electric, http://www.bmw-i-usa.com/en_us/, (checked on October 21, 2012)
8. Vilimek, R., Keinath, A., Schwalm, M.: The MINI E field study. Similarities and differences in international everyday EV driving. In: Advances in Human Aspects of Road and Rail Transportation. Advances in Human Factors and Ergonomics Series, pp. 363–372. CRC Press (2012)
9. BMW Group: BMW Delivers First All-Electric ActiveE in the US. (2012), https://www.press.bmwgroup.com/ (checked on October 28, 2012)
10. Nardi, B.A., Schiano, D.J., Gumbrecht, M., Swartz, L.: Why we blog. Communications of the ACM 47(12), 41–46 (2004)

Do All People Enjoy the Benefits from Technology Innovation?

Yoon Jeon Koh[1], Jae Heon Park[1], and Byung Do Chung[2]

[1] Economics & Management Research Lab., kt, Sungnam, Korea
{yoonjeon.koh,jaeheon.park}@kt.com
[2] Department of Systems Management Engineering, Sungkyunkwan University, Suwon, Korea
bd.chung@skku.edu

Abstract. The spread of emerging devices such as smartphone and Tablet PC enables people to experience an unprecedented technology innovation. However, some smartphone users show their intentions to reuse the feature phone due to the increased stress and pressure caused by a lot of functions of the smartphone. This study explores that the technology innovation doesn't provide all the people with the same level of values and satisfactions. A paper and pencil survey is employed with 3,798 mobile users randomly selected from 10's to 60's in Korea and asks them to answer the devices, ICT services, and applications in use, and the degree of their satisfactions as well as their perceived values and reuse intentions. Results show that the effects of technology innovation are different depending on mobile users' perceived value of technology innovation and price-value scheme, satisfaction of mobile service in use, and ages.

Keywords: technology innovation, smartphone, perceived value, satisfaction.

1 Introduction

With the development of technology, the users of the emerging devices such as smartphone and tablet PC are rapidly growing. For example, the penetration rate of mobile phones in Korea has already surpassed 100% and about 60% of mobile phone users switched feature phones to smartphones in 3 years only after kt first introduced the iPhone into the Korean market in 2009. The tablet PC market is also growing in Korea, even though the usage rate is less than 20% out of the mobile phone users now.

The spread of emerging devices mentioned above enables people to experience an unprecedented technology innovation in User Interface (UI) such as voice search and touch-sensor screen. In addition, the technology innovation enables people to use diverse functions of the new devices in a simple and easy way. As a result, people become more enthusiastic for the UI which is regarded as a key factor to influence the competitiveness of the emerging devices, driving the direction of the technology development such as motion and voice sensing.

However, some smartphone users show their intentions to reuse the feature phone due to the increased stress and pressure caused by the smartphone such as frequent

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 402–410, 2013.

email checks and notification sounds. On the other hand, a newest smartphone with high-specification is a just useless device to those who are not familiar with applications and functions which are based on the new UI, or those who use only voice and text services with their mobile phones. Thus, this study attempts to explore that the technology innovation doesn't provide all the people with the same level of benefits, values, and satisfactions, which are influenced by their life style and attitudes toward the technology.

2 Literature Review

Research has been shown that customer satisfaction is an important factor positively influencing customer loyalty and post-purchase intention in telecommunication market (Kim et al., 2004; Kuo et al., 2009). However, customer satisfaction is a broad and complex concept, which can be viewed from various perspectives including price, product design and service quality. Product design and demographic factors have been investigated as influencing the level of customer satisfaction of a mobile phone. For example, Han et al. (2004) and Ling et al. (2007) investigated critical design features of a mobile phone affecting overall user satisfaction. Yuen et al. (2011) examined Malaysian citizens to indentify the demographic factors influencing customer satisfaction.

More attentions have been given to new characteristics and functionalities due to an increasing number of value-added features on current mobile phones, especially smartphones. A recent study by J.D. Power and Associates (2011) found that the key factors of overall satisfaction with a smartphone were not only design features such as a specific size and weight but latest technological advancement. Among the studies on the relationship between the new features and customer satisfaction, Park and Lee (2011) analyzed how user experience and device characteristics of smartphone affect customer satisfaction. In their pilot study, they showed that instant connectivity is the most important factor but smartphone stress and enjoyment aren't. Ling et al. (2006) showed that color screen, mobile Internet and voice dialing strongly impacted satisfaction level. Also, Haverila (2011) found that the most important feature is battery/talk time among male users. Dobrota et al. (2012) analyzed how customer satisfaction is related to the customers' foreknowledge, awareness and information about mobile operation system. Bayraktar et al. (2012) compared customer satisfaction and loyalty efficiency of six mobile phone brands in Turkish telecommunication market.

While a lot of research has focused on the relationships between new features of smartphone and customer satisfaction, some have shown the relationships between smartphone price and customer satisfaction. In Tan et al. (2012), the most important feature of a smartphone is price plans by telecom operators rather than other features including technology-related features, design-related features, application-related features and image-related features. Also, smartphone price is ranked 5th among 11 smartphone features. Lee (2011) claimed that the price of a smartphone has no relationship with customer loyalty to the mobile service provider.

According to Kuo et al. (2009), customer's perceived value is also closely related to customer satisfaction and can be considered from different perspectives including money, quality, benefit and social psychology ((Bishop (1984), Zeithaml (1988), Sheth (1991)). From the perspective of quality in mobile services, studies have investigated the perceived value of short message service (Lai (2004)), mobile commerce (Lin and Wang (2006)), mobile value-added services (Kuo et al. (2009)). In addition, self-disclosure, flow and social presence have been considered as affecting user satisfaction of instant mobile messenger in smartphone (Park et al. (2011)). However, few have explored the relationship between the satisfaction of smartphone and the perceived value of technology innovation of smartphone. Therefore, this study will explore how users' perceived value of technology innovation of smartphone affects their satisfaction of the smartphone, and include personal characteristics like price-value perception and the satisfaction of their mobile service provider as important factors.

3 Research Model and Hypotheses

The research model includes demographic and psychological variables determining users' satisfaction of their smartphones as shown in Fig. 1. We assumed that Perceived Value of Technology Innovation (PVTI) and Price-Value Perception (PVP) would determine Satisfaction of Smartphone (SS). We also hypothesized that demographic characteristics (Gender and Age) and users' satisfaction of their mobile services (SMS) would be related to Satisfaction of Smartphone (SS).

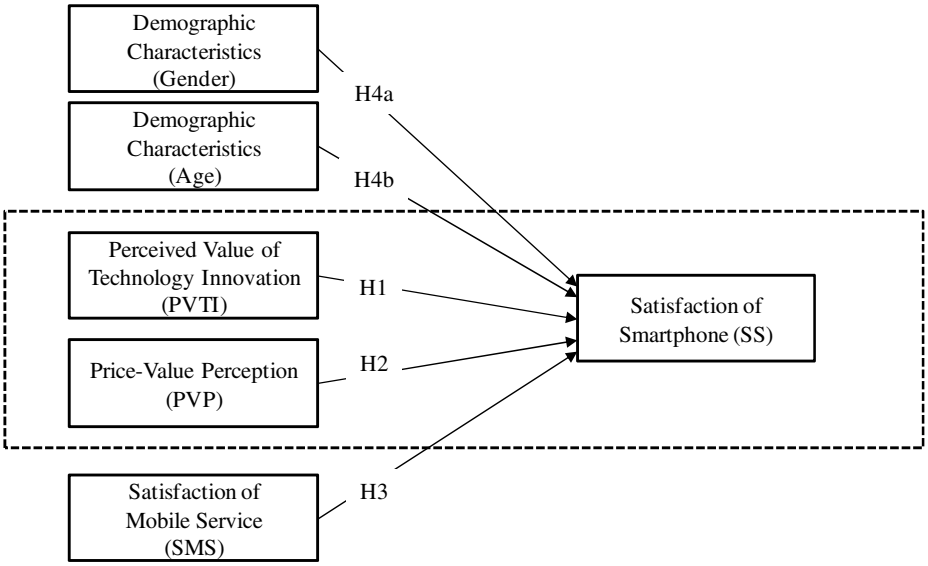


Fig. 1. Research Model

Perceived Value of Technology Innovation (PVTI) of a mobile phone represents how much a smartphone user recognizes and utilizes the innovative value of a smartphone in contrast to a feature phone. The innovative value of a smartphone can be explained by the function of data communication. Although a feature phone also provides data communication services, they are very constrained compared to a smartphone. PVTI was measured in terms of the perceived ratio of the value of data communication in the whole value of smartphone and the perceived ratio of the value of application messenger in users' communication usages. Considering that customer's perceived value is related to customer satisfaction, we propose the first hypothesis:

Hypothesis 1. PVTI is positively related to SS.

In general, smartphone users pay higher telephone rates as well as use more expensive devices, compared to feature phone users. All smartphone users do not think that higher price provides higher value or quality, affecting the satisfaction of a smartphone. We measured Price-Value Perception (PVP) in terms of user's price-quality schema (Bearden et al., 2011). Therefore, the second hypothesis is:

Hypothesis 2. PVP is positively related to SS.

Mobile phone users use mobile voice and data communication services. Because functions of a mobile phone are realized based on a mobile telecommunication service, we assumed that SMS affects SS. Thus, we hypothesize the following:

Hypothesis 3. SMS is positively related to SS.

Utilities of technology are various according to users' demographic characteristics. In general, individual's gender and age are basic considerations in segmentation for marketing. It is important to examine the effect of users' gender and age on the satisfaction of a smartphone. Thus we propose the following hypotheses:

Hypothesis 4a. Gender is related to SS.

Hypothesis 4b. Age is related to SS.

4 Methods

This study employed a face-to-face survey to obtain 3,798 questionnaires. The respondents were randomly selected from all of major cities and provinces of Korea, according to the national demographic statistics. Thus, the survey data represents a summarization of nationwide characteristics of Koreans. Smart PLS was used in the analysis of this study. First, a measurement model was tested using confirmatory factor analysis. For the test, individual item reliability, internal consistency and discriminate validity were analyzed. After the measurement model test, a structural equation model was analyzed and the hypotheses were tested.

Table 1. Characteristics of Respondents

Category		Frequency	Percentage (%)
Gender	Male	2003	52.7
	Female	1795	47.3
Age	13~19	458	12.1
	20~29	783	20.6
	30~39	905	23.8
	40~49	876	23.1
	50~59	646	16.7
	60~69	140	3.7
Smartphone Manufacturer	M1	2099	55.3
	M2	410	10.8
	M3	448	11.8
	M4	489	12.9
	Etc.	352	9.3
Mobile Service Provider	T1	1898	50.0
	T2	1191	31.4
	T3	677	17.8
	Etc.	32	0.8

5 Results

5.1 Reliability and Validity of Measures

Individual item reliability was tested using factor loading of measured variables for each latent variable. If the factor loading of an item are greater than 0.5, the item reliability is demonstrated. Composite reliability indexes were used to test internal consistency of latent variables. The acceptable level of composite reliability index is 0.7. To test the discriminant validity, the square root of AVE (Average Variance Extracted) was used. It should be greater than 0.5 to verify the discriminant validity. Table 2 shows the results of the reliability and validity test. Factor loadings are greater than 0.5 and the square root of AVE and composite reliability are greater than 0.5 and 0.7, respectively. Thus, reliability of measured variables and discriminant validity of latent variables are justified.

5.2 Analysis of the Structural Model

Bootstrapping of Smart PLS was used to analyze the research model. Table 3 and Fig. 2 show the results. PVTI and PVP were positively related to users' satisfaction of smartphone (SS). The path coefficients of them are 0.100 and 0.091, respectively, and p-values are less than 0.001. Thus, H1 and H2 are supported. SMS is found to be significantly related to SS. The path coefficient is 0.324 and p-value is less than 0.001. H3 is supported. The path between Gender and SS was found to be insignificant. The p-value is greater than 0.05 and thus H4a is not supported. On the other hand, Age is related to SS. The path coefficient is 0.038 and p-value is less than 0.05.

Table 2. Results of Validity and Reliability Test

	Factor Loading	SQRT (AVE)	Composite Reliability
Perceived Value of Technology Innovation		0.86	0.85
What do you think is the ratio of the value of mobile internet service among all of the services provided by mobile communication?	0.87		
What do you think is the ratio of the usage of communication though messenger application among all of your mobile communication?	0.86		
Price-Value Perception		0.94	0.96
Generally speaking, the higher the price of the product, the higher the quality	0.94		
The price of a product is a good indicator of its quality	0.94		
You always have to pay a bit more for the best	0.93		
Satisfaction of Mobile Service		1.00	1.00
How much are you satisfied with your mobile communication service provider?	1.00		
Gender		1.00	1.00
What is your gender?	1.00		
Age		1.00	1.00
What is your age?	1.00		
Satisfaction of Smartphone		1.00	1.00
How much are you satisfied with your smartphone?	1.00		

Table 3. Results of Hypotheses Test

Hypothesis	Path	Path Coefficient	t-value	p-value
H1	Perceived Value of Technology Innovation → Satisfaction of Smartphone	0.100	5.305	0.000
H2	Price Value Perception → Satisfaction of Smartphone	0.091	4.966	0.000
H3	Satisfaction of Mobile Service → Satisfaction of Smartphone	0.324	18.508	0.000
H4a	Gender → Satisfaction of Smartphone	-0.018	1.200	0.237
H4b	Age → Satisfaction of Smartphone	0.038	2.097	0.042

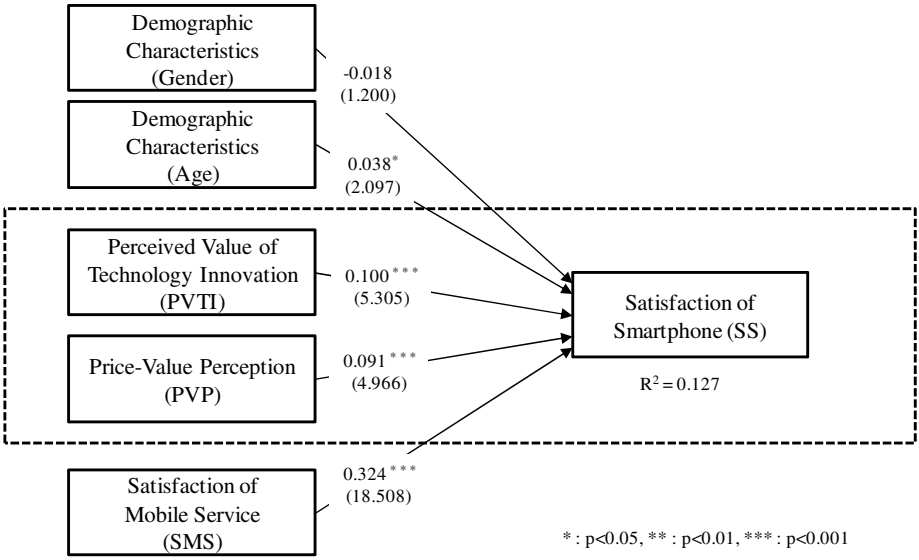


Fig. 2. Results of Research Model and Hypotheses Test

6 Discussions

6.1 Summary of Findings

This study provides a strong evidence that the satisfaction of smartphone is affected by the perceived value of technology innovation of a smartphone. Two measured variables, the value of mobile internet service and the ratio of the usage of communication through messenger application, were used and showed to be reliable to represent the perceived value. This means that although people accept new technology, they have different acceptance level of value provided by the technology and thus the satisfaction of a new technology is influenced by not only its features but also users' acceptance level.

Another finding is that the relationship between the price-value perception and the satisfaction of smartphone is significant. This means that some people may be dissatisfied by the high price although they get great value from technology innovation. A user's satisfaction with the mobile service provider he/she subscribes to is a significantly associated with the satisfaction of the smartphone. This shows that satisfaction of a technology which is enabled by another one is influenced by the other one. As expected, a user's age has a significant relation with the satisfaction of smartphone. On the other hand, the gender is not significantly related with the satisfaction of a smartphone. This shows that there are some other factors which are affecting on the satisfaction of a smartphone and have in relation with age but not with gender.

6.2 Contributions and Limitations

This study expands the concept of perceived value by showing that people show the perceived value from the perspective of technology, in addition to that of money, quality, and benefit. The results show that the higher the perceive value of a smartphone is, the higher the satisfaction of smartphone is. Additionally, the findings suggest that individuals' price-value perception affects the satisfaction of smartphone.

This study also provides that the satisfaction with a telecom service provider affects the satisfaction of a smartphone, meaning that the satisfaction of a smartphone should be considered by the combination of the device and the service of a smartphone. Finally, findings suppose that all people do not fully enjoy the benefits from the technology innovation, showing that users have different acceptance levels of the smartphone value. Thus, different approaches and services should be prepared depending on the level of acceptance, in order to help people who do not easily follow and accept the technology innovation.

In spite of those contributions, this study has some limitations. Although the study supposes that the users' personal characteristics have a relation on the satisfaction of a smartphone, other factors were not considered except the price-value perception. Such as involvement and technology readiness can also be used as affecting factors. The model can be expanded to include users' intention to repurchase a smartphone to build a model explaining the relations from the perceived value to retention through satisfaction of technology innovation. These limitations should be our further works.

References

1. Kuo, Y.F., Wu, C.M., Deng, W.J.: The relationships among service quality, perceived value, customer satisfaction, and post-purchase intention in mobile value-added services. Computers in Human Behavior 25(4), 887–896 (2009)
2. Kim, M.K., Park, M.C., Jeong, D.H.: The effects of customer satisfaction and switching barrier on customer loyalty in Korean mobile telecommunication services. Telecommunications Policy 28(2), 145–159 (2004)
3. Han, S.H., Kim, K.J., Yun, M.H., Hong, S.W., Kim, J.: Identifying mobile phone design features critical to user satisfaction. Human Factors and Ergonomics in Manufacturing & Service Industries 14(1), 15–29 (2004)
4. Park, Y., Chen, J.V.: Acceptance and adoption of the innovative use of smartphone. Industrial Management & Data Systems 107(9), 1349–1365 (2007)
5. Ling, C., Hwang, W., Salvendy, G.: Diversified users' satisfaction with advanced mobile phone features. Universal Access in the Information Society 5(2), 239–249 (2006)
6. Yuen, Y.Y., Yeow, P.H., Connolly, R.: Mobile phone satisfaction in Malaysia: a demographic analysis. International Journal of Mobile Communications 9(4), 341–358 (2011)
7. Ling, C., Hwang, W., Salvendy, G.: A survey of what customers want in a cell phone design. Behaviour & Information Technology 26(2), 149–163 (2007)
8. Dobrota, M., Milenković, I., Bulajić, M.: The Influence of Foreknowledge on Customers'Satisfaction With Mobile Operating Systems. Business Intelligence and Decision Making in Management 251

9. Bayraktar, E., Tatoglu, E., Turkyilmaz, A., Delen, D., Zaim, S.: Measuring the efficiency of customer satisfaction and loyalty for mobile phone brands with DEA. Expert Systems with Applications 39(1), 99–106 (2012)
10. Bishop Jr., W.R.: Competitive intelligence. Progressive Grocer 63(3), 19–20 (1984)
11. Zeithaml, V.A.: Consumer perceptions of price, quality, and value: a means-end model and synthesis of evidence. The Journal of Marketing, 2–22 (1988)
12. Sheth, J.N., Newman, B.I., Gross, B.L.: Consumption values and market choices: Theory and applications, pp. 16–74. South-Western Pub. (1991)
13. Lai, T.L.: Service quality and perceived value's impact on satisfaction, intention and usage of short message service (SMS). Information Systems Frontiers 6(4), 353–368 (2004)
14. Lin, H.H., Wang, Y.S.: An examination of the determinants of customer loyalty in mobile commerce contexts. Information & Management 43(3), 271–282 (2006)
15. Park, B.W., Lee, K.C.: A Pilot Study to Analyze the Effects of User Experience and Device Characteristics on the Customer Satisfaction of Smartphone Users. Ubiquitous Computing and Multimedia Applications, 421–427 (2011)
16. Park, S., Oh, D., Lee, B.G.: Analyzing user satisfaction factors for instant messenger-based mobile SNS. In: Park, J.J., Yang, L.T., Lee, C. (eds.) FutureTech 2011, Part II. CCIS, vol. 185, pp. 280–287. Springer, Heidelberg (2011)
17. Lee, J.W.: Critical factors promoting customer loyalty to smartphone and mobile communications service providers (2011)
18. Bearden, W.O., Netemeyer, R.G., Haws, K.L.: Handbook of Marketing Scales, 3rd edn., p. 379. SAGE Publications (2011)
19. Tan, W.K., Yi-Der Yeh, S.J.C., Lin, Y.C., Kuo, C.Y.: How Consumers Assess Product's Features?: A Case Study of Product Features of Smartphone. In: Proceedings of the 6th International Conference on Applied Mathematics, Simulation, Modelling, pp. 131–137. World Scientific and Engineering Academy and Society, WSEAS (March 2012)
20. Power, J.D.: The Right Blend of Design and Technology is Critical to Creating an Exceptional User Experience with Smartphones and Traditional Mobile Devices. JD Power and Associates (2011)

Semantically Integrated Business Applications for Enterprise Resource Planning Systems

Patricia Kraft and Rainer Thome

Julius-Maximilian-University, Chair of Business Information Systems,
Wuerzburg, Germany
{pkraft,thome}@wiinf.uni-wuerzburg.de

Abstract. The development of enterprise resource planning (ERP) software or enterprise information systems (EIS) has always been closely linked to the discussion on integration. Over time, integration of software solutions has improved considerably. This has been achieved through several changes made in the past, but mostly it has become possible thanks to cutting-edge technologies. A new approach is semantically integrated business applications that enable the evolution of standard end-to-end software.

Keywords: Business applications, software development, enterprise resource planning (ERP), integration, add-on, enterprise information systems.

1 Introduction

The development of solutions for supporting the processes and organization of companies with enterprise resource planning (ERP) software has always been accompanied by a call for more integration. The ERP systems are the target of a great many demands, all of which various advocates expect the systems to meet. Depending on the specific point-of-view, some demands seem more pressing than others. These include calls for more flexibility, because in adapting to the real world, organizations are subject to constant change. But software should support change rather than be so inflexible as to hinder it and reduce an organization's competitiveness. Costs are also an issue. Enterprise software is not a primary driver of profit. As a result, companies strive to keep the cost of their information system as low as possible. A third point of conflict is functionality. Organizations' user departments raise this point because to them, the system is only beneficial if it can model all relevant corporate processes. But additional functionality increases complexity, compromising a system's transparency and rendering it inflexible. What is more, functionality is the main cost driver. Integration receives the weakest representation in this discussion. It is often assigned the lowest priority because its advocates are few and have little influence. In an ERP system, integration (of technical aspects and of applications) is the key to the smooth operation and modeling of processes, decreased maintenance costs and consistent data pools. Integration can also include

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 411–417, 2013.

the look and feel of the user interface, providing consistency for the user. But these points are not concrete, cannot be implemented immediately, are difficult to quantify, are not a priority for management and are therefore at a disadvantage when compared to the interests previously mentioned.

This paper will aim to answer the question: What changes will cloud-based solutions bring to ERP software and how will these changes influence integration of software development long-term? First, an overview of the research done on integration of ERP solutions will be presented and the most recent developments in cloud-based ERP systems described. Next, semantically integrated applications, an integrated method for developing business application software, will be introduced.

2 State of the Art

FINK and HANSEN/NEUMANN define the term enterprise resource planning (ERP) system as a multiple-component business application system for end-to-end information processing. These systems are used to model and execute processes in all user departments of an organization (e.g. Production, Procurement, Accounting, etc.). In addition to other manufacturers, such as Oracle and Microsoft, SAP is the worldwide market leader for business application software both for large companies and for small and mid-size businesses [1][2].

Development in the area of *integration of business application software* up to the present can be divided into phases or blocks along a timeline. HANSEN and NEUMANN created a breakdown (Table 1) that divides development into five distinct phases. These five phases reveal development toward increasingly open platforms, which can communicate with each other via the Internet and provide a network that extends beyond an organization's borders. In 2005, when this hypothesis was posed, cloud solutions were not prevalent for ERP software. This thought has only begun to push its way into the conversation since a range of offerings in this sector has been launched onto the market and gained widespread popularity.

Software as a service can be described as a software delivery and business model by which the software vendor enables customers use of the standard software applications via the Internet as a service. In contrast to traditional on-premises software applications, the customer does not acquire a software license, but subscribes to the service. The software itself is hosted by the software vendor and does not require implementation or deployment at the customer site. The specific pricing model for SaaS is commonly based on a usage metric, such as per user per month." [3]. Solutions also exist as part of the platform as a service model. In this case, the object of examination is not just software but a platform that unites multiple offerings and vendors selling their cloud solutions and services. In other words, the platform is actually the model of a partner ecosystem that delivers additional offerings for the solutions available [4].

Table 1. Integration phases of enterprise software [2]

	Phase 1 1960-1980	**Phase 2 1980-1990**	**Phase 3 1990-2000**	**Phase 4 2000-2010**	**Phase 5 2010-2020**
Area	Isolated functions	Application areas	Internal business processes	Cross-company processes	Extensive corporate networking
IT technology	Customer-specific programming, classic data organization for proprietary platform	One manufacturer's proprietary application software on a few proprietary platforms	One manufacturer's proprietary application software on many proprietary platforms	One manufacturer's proprietary application software on many proprietary or open platforms	Combination of several manufacturers' application software on open-source platforms
	Terminals connected to mainframe computers	Terminals connected to mainframe computers	Client-server system	Client-server system, Internet-based	SOA system, web services
	Text-oriented UI	Text-oriented UI	Graphic UI	Web-based UI (browsers)	Web-based UI (browsers)

Up to now, the prevailing systems on the ERP market have been on-premises solutions. The customer installs, maintains and adds to these, if more functionality is desired. These add-ons are normally intended for a single customer, since each customer has his or her own requirements. Add-ons are either programmed by the customers themselves, or customers hire a software manufacturer to develop a solution. This gives rise to a number of problems: it compromises future ERP software releases, on the one hand, and on the other, customer-specific programming of the solution causes maintenance costs to skyrocket. All changes to the ERP system that go beyond those made via customization settings (adjustment of parameters) hamper flexibility and complicate maintainability. This knowledge leads to the demand for standard software and the adjustment of organization and processes to the ERP solution within its scope of performance. The continuation of this thought results in Continuous System Engineering (CSE) by THOME, the ongoing development of application software in interaction with the changes in corporate organization [5].

In the past decade, the issue of integration has been examined by a number of authors but never defined completely and conclusively. Many specialists in the field have attempted to quantify integration, and have illuminated it from various angles. MERTENS, for example, makes a distinction between the range, direction, object and scope of integration [6]. HAHN advocates a different approach. He distinguishes among data, model, and process integration. These levels can be closely or loosely

linked to one another [7]. JANSEN and HUFGARD introduce the business aspect. HUFGARD et al. maintain that the term business integration includes technical aspects of information processing, as well as organizational and marketing aspects [8]. JANSEN's "six target levels of integration" comprise strategic, organizational, human resources, cultural, operative and external levels of integration [9]. This diversity of definitions makes clear the complexity of the term integration and its far-reaching influence on an organization's processes. In sum, one can say that for integrated business processes, it is absolutely essential to be able to:

- access all relevant data
- save it without redundancy immediately after its creation,
- update software to state-of-the-art levels, and
- communicate digitally with other business partners across systems and company borders.

The user may not be impacted by changes in the system or in the look and feel of the user interface and yet, adjustments made necessary by organizational or business-related changes must be integrated into the system quickly and seamlessly.

3 Cloud ERP Systems Today

When comparing simple, cloud-based solutions to their locally installed competitors, one might wonder whether saving data and applications on external servers makes a difference with respect to integration. A cursory glance can lead to suspicions that cloud-based solutions are merely a marketing campaign and that the flashy new façade of cloud products conceal only externally-hosted variations of existing solutions. At first glance, the disadvantage of saving data in the cloud is that customers lose data sovereignty over their most valuable corporate information. But this argument can be refuted by the simple reflection that most organizations that purchase cloud applications are small and midsized businesses that do not have the necessary resources to operate their own high-security data centers. In this case, the data is probably safer in the hands of an external service provider than on the company's own premises. Another question has to do with workflow. When a web browser is used, it might seem as though users are more remote than when they use an in-house solution. But really, accessing an ERP system via the Internet is simpler and more secure than accessing an internal solution that is purposely blocked off from the external world and can be accessed only remotely via secure VPN entry points, if at all. The result is that, for example, sales reps can better access a cloud solution from a customer's premises than access a solution installed solely within their own company. SaaS improves networking and availability of all corporate data.

Regardless of the availability of and security issues concerning these solutions, the software structure holds the key to groundbreaking changes. Popular solutions on the market, such as NetSuite and Salesforce.com all offer the benefits of a web application, but behind their façades, they are nothing more than conventional software packages, i.e. software as a service [10]. But SAP's new Business ByDesign solution has a completely different structure; it is a whole new type of application.

At the heart of SAP Business ByDesign is Business Configuration, a catalog for setting all parameters of the ERP solution. It forms the basis for the application's configuration, the modeling of each customer's requirements. Users answer simple Yes or No questions to create, check and implement organization characteristics relating to system functionality. And then, depending on the answers to the questions, the application areas are parameterized with the help of a fine-tuning tool. It recognizes soft factors in the organization, migrates data and creates structures. This includes, for example, defining various models for employee working hours and migrating customer master data via an XML interface [11] [4].

Another unique component of ByDesign is the cloud-based development environment, Business ByDesign Studio. It can be used to develop add-ons that go beyond what is offered in the standard solution, to set up a link to external software, and to customize the standard software components. These two types of functionality differ very little from capabilities available in on-premises solutions. Customer-specific programming makes it possible to create each ERP solution in line with customer needs and to add functions. However, making changes in ByDesign via public solution models (PSM) is a completely different story. PSM contains a collection of data fields, methods and structures that programmers (even external programmers) can access via the ByDesign Studio development environment, in order to program additional functionality. PSM content adapts to future releases, which ensures that a one-off solution will still work after a release upgrade. Whenever a partner company develops an add-on for Business ByDesign, it automatically becomes part of the overall solution. The new components are sold in the SAP Store (a platform like an app store that offers add-on solutions as applications). When customers purchase a solution, it appears in their business configuration and they can make immediate use of it [4].

4 Infrastructure for Business Application Development and Integration of Applications

The approach just described for developing add-ons for a cloud-based solution – fusing add-ons with the host's infrastructure and enabling them to become part of the overall solution – offers a higher degree of integration from the perspective of all stakeholders than the approach previously used to develop add-ons for on-premises solutions. For the developer, using SAP software development kits (SDK) closely links new functions to the overall solution. The development environment is based on the PSM settings and contains specific patterns for creating the user interface that correspond to the solution and its outward appearance. The user cannot tell which screen is part of the standard solution and which has been added. The development environment ensures that all requirements are met for an integrated, compatible solution that will also work with future release upgrades.

For the customer, this approach improves business flexibility; the system offers multiple setting options, and add-ons from the SAP Store can easily be downloaded to it. The vast number of customers it benefits also makes it less expensive than company-specific developments for on-premises solutions. And, for the first time it has

made possible an objective comparison of alternative business applications on the market and their respective providers. If a solution is not available, a customer can ask an SAP partner company to create it. If a solution becomes obsolete because, for example, an organization has changed its structure, the user can very simply deactivate the solution. As a result, an organization is not weighed down with superfluous coding from one release to the next.

The provider and platform hosts face higher infrastructure costs by providing the solution, and lower maintenance costs with respect to new releases. "The SaaS vendor provides multiple tenants, one tenant per customer, based on a shared technology stack including hardware infrastructure, database and application" [3]. Integration of infrastructure is seamless because the tenants share the same hardware. Because all users are forced to upgrade to the new release during the two-to-three-month transition phase, a maximum of two releases are maintained – the current and the future release. The entire infrastructure is state-of-the-art and no legacy systems need to be maintained. All stakeholders are in the same boat. Web services support communication with external programs. They unite the communications structures and ensure semantically-correct transmission of data. Communication between various different business partners, all with separate ERP systems, is done this way, ensuring seamless communication across company and system borders.

If one compares this to the degree to which ERP systems are integrated today, marked discrepancies become apparent. The transition from the on-premises world to that of cloud-based solutions with the option to add semantically integrated business applications is comparable to the step from terminal computers to desktop computers. A stand-alone, partially linked solution becomes a completely integrated standard solution that grows in proportion to user demands.

In sum, semantically integrated business applications offer the following advantages:

- Development takes place within an SDK provided by the manufacturer,
- The user cannot distinguish between the standard solution and the add-on, with respect to the look and feel of the user interface,
- The add-on is an integral part of the functionality within the basic ERP solution,
- The new function does not compromise future upgrades for the overall solution,
- The solution is activated and configured along with the default functions (e.g. via Business Configuration),
- All customers who have the default system can use the solution,
- The solution is sold at a marketplace that is equally accessible to all customers.

Systems that meet all of the above-mentioned criteria are considered semantically integrated business applications for cloud-based ERP solutions.

References

1. Fink, A., et al.: Grundlagen der Wirtschaftsinformatik, p. 207. Physica, Heidelberg (2005)
2. Hansen, H., Neumann, G.: Wirtschaftsinformatik 1, p. 529. Lucius & Lucius, Stuttgart (2005)

3. Bandulet, F., et al.: Software-as-a-Service as Disruptive Innovation in the Enterprise Application Market. In: Benlian, A., et al. (eds.) Software-as-a-Service. Anbieterstrategien, Kundenbedürfnisse und Wertschöftungsstrukturen, pp. 16–17. Gabler, Wiesbaden (2010)
4. Schneider, T.: SAP Business ByDesign Studio. Application Development, pp. 24–28. Galileo Press, Boston (2012)
5. Thome, R.: Grundzüge der Wirtschaftsinformatik, pp. 149–159. Pearson, Munich (2006)
6. Mertens, P.: Integrierte Informationsverarbeitung 1 - Operative Systeme in der Industrie, p. 2. Gabler, Wiesbaden (2007)
7. Hahn, A.: Integrationsumgebung für verteilte objektorientierte Ingenieursysteme, pp. 7–23. Universität Paderborn Heinz Nixdorf Inst., Paderborn (1998)
8. Hufgard, A., et al.: Business Integration mit SAP Lösungen, pp. 15–16. Springer, Berlin (2005)
9. Jansen, S.: Mergers & Aquisitions: Unternehmensakquisitionen und -kooperationen. Eine strategische, organisatorische und kapitalmarkttheoretische Einführung, p. 330. Gabler, Wiesbaden (2008)
10. Salesforce: Salesforce.com., `http://www.salesforce.com/company/`
11. Hufgard, A., Krueger, S.: SAP Business ByDesign: Geschäftsprozesse, Technologie und Implementierung anschaulich erklärt, pp. 154–177. SAP Press, Bonn (2012)

Towards a Next Generation Universally Accessible 'Online Shopping-for-Apparel' System

Kasper Kristensen, Nanna Borum, Line G. Christensen, Henrik W. Jepsen, Jacob Lam, Anthony L. Brooks, and Eva Petersson Brooks

Aalborg University Esbjerg, Niels Bohrs vej 8, Esbjerg, Denmark
{tb,ep}@create.aau.dk

Abstract. This paper reports initial research findings from on-going longitudinal participatory design studies within a national (Danish) funded project to realize a gesture-controlled 'Online shopping-for-apparel' system – A Virtual Dressing Room (VDR). A product that reduces customer purchase returns, reportedly up to 40%, which is a huge burden to the clothing industries as shopping percentile of sales online continues to increase, is targeted. Three studies are reported where results cumulate to highlight the need for continued research to realize a next-generation system to improve the user experience of online shopping for apparel where conclusions point to the need for adaptive user interface improvements. Unforeseen was that wheelchair-bound public especially responded positively to the potentials for the concept due to their limited mobility in shopping and this accessibility aspect can be a significant future market.

Keywords: Online shopping system, purchase returns, e-shopping experiences.

1 Introduction

The United Kingdom Office of National Statistics reveals that the average weekly value for Internet retail sales in June 2012 was estimated to be £493.3 million[1]. It is predicted that by 2018, 35% of clothing sales will be via the Internet[2]. However, current figures show that between 15-40% of apparel purchased online (e-shopping) is returned because customers deem they do not fit[3] or do not look right[4]. Trust policies that safeguard customer purchases enabling such returns are influential and, in some cases, these policies are reported as being abused whereby a purchase is received, worn, and then returned with full credit or refund. In the clothing industry, such high volume of returned apparel is economically disastrous and it is in respect of this

[1] http://www.ons.gov.uk/ons/rel/rsi/retail-sales/june-2012/stb-june-2012.html

[2] Heikki Haldre, a founder of Fits.me, an online fitting room service that creates custom virtual mannequins for people to dress with clothes before buying them.

[3] http://online.wsj.com/article/SB10001424052702304724404577293593210807790.html

[4] http://www.imrg.org

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 418–427, 2013.

problem that the project is directed. Thus, the main goal of this work is to realize an online shopping-for-apparel system in the form of a Virtual Dressing Room (VDR) where contemporary camera-based ICT is used to reduce consumer returns.

In order to develop a turnkey solution, a participatory design (PD) approach has been applied in line with Brandt [1]. Fundamentally PD represents a set of theories, practices and studies whereby user communities play a substantive role in activities that can lead to the creation of software and hardware computer technologies and their application in real-world contexts [2-4]. The benefits of adopting such an approach for the design of the VDR system include a better understanding of the reasoning behind online purchasing behavior. This is important particularly given that these numbers are expected to rise even more in the future with predictions of online sales in Western Europe to increase at a 10% compound annual growth rate over the next 5 years[5]. Findings from the PD field studies are periodically fed into the technical partners process pipeline to support development decisions. This paper presents the result from the two initial phases of the design process, namely the preparation and the incubation phases (c.f. Seifert et al. [5]), which represent the initial inspirational and definition phases where the problem was defined and potential solutions identified.

2 Related Work

In the clothing industry many fit technologies have been tried and flopped. For example, retailers have created virtual mannequins for customers to dress and set up full-body scans in stores. It was found that shoppers are reluctant to use systems that require much effort and time unless a special purchase is the goal. It is acknowledged that for women's clothing in particular, sizing is difficult for many complex issues including the psychological. One brand's size 12 is another's 10 or 14, primarily because fashion labels shape their clothes so differently, using their own closely guarded specifications to create patterns. Online retailers often ask shoppers to consult "fit charts" or type in body measurements, which can be time-consuming. However, the True Fit system reports increased sales and reduced returns up to 30% (premier denims). Consumers create a profile of age, height, weight and body shape. Then the customers select items that fit well from their own closets and identify the brands, styles and sizes to True Fit. Another personalized online shopping approach by Stitch Fix[6] involves female customers filling out a similar profile form whereafter online photos of clothing and accessories are rated to give the company an idea of the client's taste. NoMoJeans[7] takes precise customer measurements with a 3D body scanner which are kept on a data base for future purchases.

This paper reports initial trendspotting (field) and participatory design (PD) studies, conducted to unfold tendencies, key patterns and personas related to on- and offline shopping. The trendspotting included structured interviews with different user

[5] http://www.forrester.com
[6] https://www.stitchfix.com
[7] http://www.nomojeans.com

segments representing consumers and retailers/industry. The consumer segment were women 18-30 years of age (n=15). Furthermore, observations of consumer shopping behavior was conducted in combinations with focus group interviews with females 18-35 years of age (n=7), expert interviews (sales staff) (n=3), and document analysis (magazines, shopping forums, statistics, industry reports).

A field study was conducted at a public trade fair in Copenhagen, which included video observations, questionnaires and unstructured interviews (n=11) where the freely available LazyLazy[8] system was used. Building on these studies a PD study was undertaken at a shopping mall in Denmark's second city Aarhus. This included video observations, questionnaires and a focus group (n=31; females n=20). Following, a low fidelity prototyping approach was used in order to get useful design feedback from students attending Aalborg University Esbjerg, Denmark.

3 Retailer/Industry Trend Spotting/State-of-the-Art Study

Five selected companies were visited to determine state of the art in the industries associated with apparel shopping in order to supplement the PD research. These were:

1. ***Inition*** - http://inition.co.uk - a preliminary prototype was presented based on a single Kinect live mirroring the customer (similar to LazyLazy). The system was demonstrated in combination with interviewing of the system creator and VP sales. Clothes are selected from an image menu and a superimposed overlay positioned onto the live video feed, thus, in this system there is no need to disrobe.
2. ***Shape Analysis*** - http://shapeanalysis.com - a proprietary multi-camera system (4) marketed in partnership with [TC]2 in USA. The customer needs to disrobe for precision scanning. Their latest systems offer a choice between a single Kinect (2D) or four Kinect system (3D – 360 degrees). The software generates an avatar (wireframe but can superimpose own face or selected clothing) giving precise measurements. The system was demonstrated by creating the researcher's body scan. Furthermore, interviews with the creator and marketing person were conducted.
3. ***Bodymetrics*** - http://www.bodymetrics.com - a proprietary multi-camera (4) system with installations in leading stores in London (Selfridges – where tested/interviewed staff and clients) and USA (Bloomingdales). The latest system uses single or multiple Kinect cameras for boutiques and home uses. The software generates an avatar that mirrors customer rotation in real-time. The interviewer viewed the system together with the CEO/creator's representative and VP sales. Furthermore, staff (2) and customers (2) were interviewed. Customers complained of a need to initilly disrobe for scanning, then having to dress to get clothing from staff in the public area, and then to go to a fitting room to disrobe again to test apparel feel and fit. This process is being changed using a female only area.
4. ***Holition*** - http://www.holition.com – as (1)

[8] http://www.lazylazy.com

5. ***Cristina Holm, Personal Image Stylist*** - http://www.cristinaholm.com - a four-camera scanner setup with software generating an avatar (static generated for precise measures – as 2). The [TC]2[9] NX16 scanner is located in a private studio where made-to-measure exclusive clothing is designed and sold. [TC]2 company launched a more affordable Kinect-based solution in February 2012[10]. The system was demonstrated with resultant scan obtained. The owner/stylist, was interviewed alongside the system inventor. The 3D avatar can be rotated by a mouse and personalized with an image of the client's face.

The protype by Holiton/Inition was without accurate measures and used flat screen monitors in landscape mode. All of the companies are developing a single Kinect for home use and multiple systems targeting boutiques, custom tailors/designers, etc.

3.1 Customer Trend Spotting Study

The customer trend spotting study resulted in a shopping behavior analysis and personas in terms of fictional characters. The shopping behavior analysis identified the following main trends:

- *Fusion between online and offline shopping*. The customers would like to see big screens in the stores where it is possible to link to the online webshop.
- *Service level*. Some of the customers considered online shopping convenient and stressed that offline service should be improved to give value to offline shopping.
- *Convenience*. The customers preferred online shopping due to that it is not bound to time and space so they will, thereby, save time; they do not need any transportation; and they do not need to wait in any line.
- *Feel, touch and realism*. The customers stressed that the feeling of the clothes, the touch, and the realistic trying-on experience are desired features that enhance the shopping.
- *Online tools enhancing offline shopping experiences*. The customers suggested that online tools, such as size and fit information, texture detail, matching of clothes to the individual's specific body type would enhance the offline shopping experience. Furthermore, social media, such as FaceBook, where the customer can share photos of the potential purchase and receive immediate feedback, would also enhance the offine shopping.

Furthermore, the customer trendspotting study showed that personal issues such as sizes in not purely a technical question for virtual dressing room solutions, it constitutes a highly emotional issue. Some of the customer suggested that this issue should be handled with humor, for example the body could be replaced with sketchy models or 3D avatars attached with humor-directed comments, e.g. "yes, you fit into this model, but you cannot breath....".

[9] http://www.tc2.com
[10] http://www.tc2.com/newsletter/2011/113011.html

Four personas, fictional characters, emerged from the customer trendspotting study:

- *The practical non-shopper*, represented by Anna, a teacher, 29 years of age. She works roughly 37 hours per week. Buys clothes one (1) to two times per year, mostly in stores. She always checks the quality in a store before ordering clothes online. She buys for practical reasons, mainly out of necessity. She can go shopping by herself but also with friends to have a good day out. *I like to try on the clothes and I like the social experience with my friends. I would definitely also buy clothes online if I could afford it. Since I cannot afford it very often, I choose to make special occasions out of the real shopping experience*.
- *The lone shopper*, represented by Jacob, university student, 23 years of age. Jacob works as a weekend waiter and buys clothes in stores and online. He usually goes for offers and mainly shops alone as he finds it a bit boring. *I mostly need functional clothing and in the end it is me who should feel comfortable wearing the clothes. Others should not decide what clothes I should wear.*
- *The social shopper*, represented by Tina, senior college student, 20 years of age. Tina works one or two days per month in a bookstore. She lives with her parents and prefers to buy clothes in stores and often with her friends. She uses 1000 DKK per month on clothes (on average). She looks for the brands and good offers when shopping. Tina shops on average 12-20 times per year and often this is combined with a day with her friends. *It is just one of those things that is more fun to do together with friends combined with café-visits.*
- The *spoiling-mom shopper*, represented by Linda, physiotherapist running own clinic, 44 years of age. Linda works roughly 37 hours per week having two children. She most often buys clothes for others than for herself. When she buys clothes for herself, it is moslty for practical reasons. Linda enjoys buying clothes in offline shops, but occasionally buys online. When buying for others, she is very aware of fashion and trends. *When I buy clothes for others, I do so because I see a style in the stores which I think would fit the person, then the price doesn't matter that much*.

3.2 Field Study

Shortly after the trendspotting study focusing on industry and costomers (i.e. sections 3.1 and 3.2) an initial field study with public interviewees was conducted at a major Scandinavian Trade Fair event that attracted more than 7000 attendees[11] to the *Health & Rehab Scandinavia* at the Bella Center, Copenhagen 22-24 May 2012.
Questioning of attendee trends in shopping (online/offline) was conducted from a large stand where videos of current state-of-the-art systems and simulations were used to question public opinions of online shopping for apparel systems (figure 1). Wheelchair bound interviewees made clear their need for improved means to purchase clothes without having to leave their homes. The perceived ease of use (PEOU) and perceived usefulness (PU) toward enhancement and reduced effort of an online

[11] http://www.liftup.dk/index.php?id=130,247,0,0,1,0

apparel purchasing system were seen as key factors for wheelchair-bound consumers. However, it is clear that the UI needs to be adaptive to overcome challenges such as illustrated in figures 2 and 3 that highlight interface/setup design considerations.

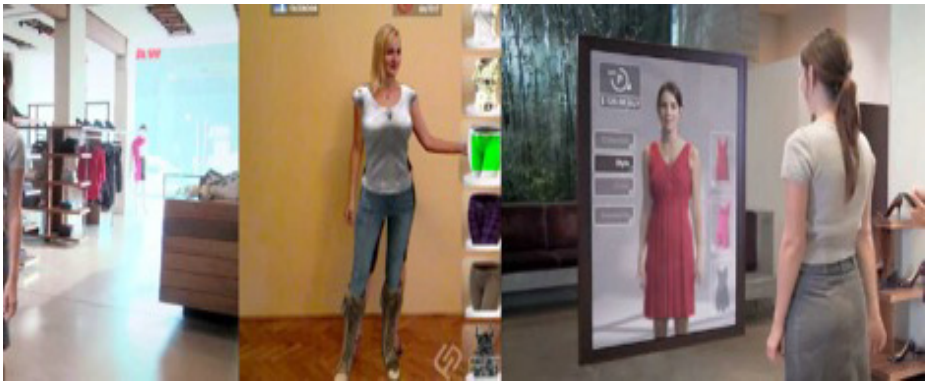


Fig. 1. Different state-of-the-art virtual dressing room solutions questioned

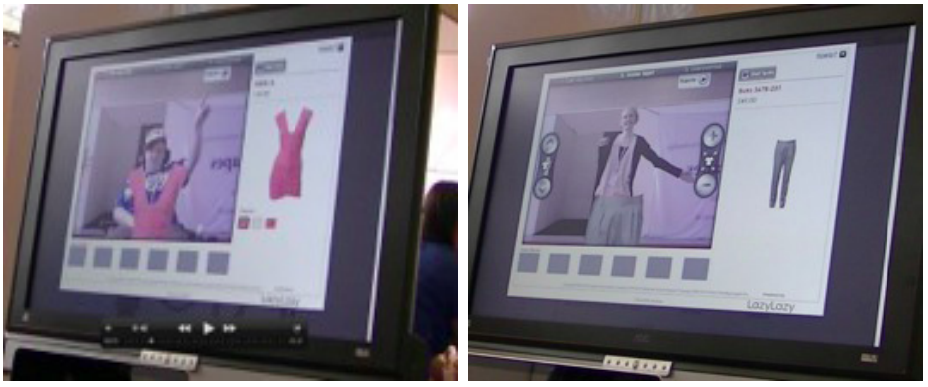


Fig. 2. Current UI problems (left; height/handedness) (right; distance view/operation)

Figure 2 (left) illustrates a participant who was in a wheelchair, which placed him lower in the picture than other participants. As a result, he could not reach the button at the top. He used solely his right hand to make selections, which provided difficulties when trying to reach the buttons on the left side of the screen. Younger children had similar issues with reaching as they were also positioned lower in the picture (due to their height). The figure 2 right image illustrates a participant who was standing too far away from the camera, which was necessary in an attempt to see her lower body and the selected to preview trousers. As a result, she had trouble reaching the buttons at the sides of the screen and the image was too small because of the distance. Figure 3 (left) illustrates how the participant had difficulties interacting with the system because the buttons were activated by motion (of other people) in the background. The right image (figure 3) illustrates persistent single-handed interaction and body twist to operate, even though participants complained it uncomfortable.

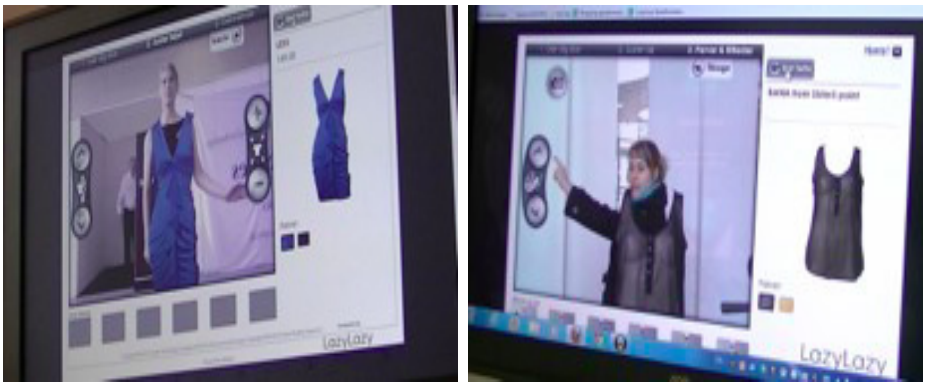


Fig. 3. Current UI problems (left; background triggering) (right; single handedness)

3.3 Participatory Design Studies

Following on from the above-mentioned investigations, two teams of university students conducted participatory design studies at a leading shopping mall in Aarhus, Denmark's second city. An online retail boutique system called LazyLazy was demonstrated to offer research observations of an online apparel system use. The data from this study led to UI investigation via focus groups of a created test image of possible improvements to the commercial product's user interface.

LazyLazy is an online shopping mall where customers can purchase via the Internet (screen capture examples figures 2 & 3). The observed demonstration set up for the studies included a flatscreen TV with a camera used by the software, and two cameras for ethnographic analysis, i.e. the prototype sessions, video observations,

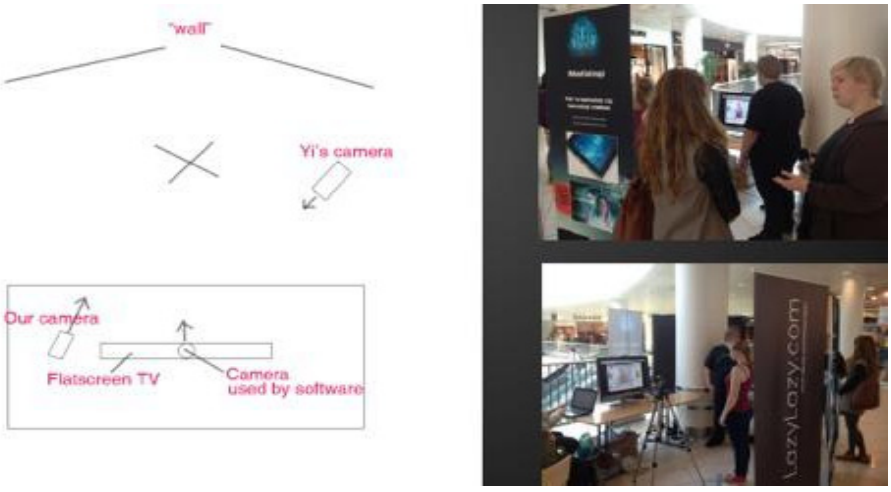


Fig. 4. The general set up of the user study at Bruun's Galleri, October 2012

questionnaires, and unstructured interviews (figure 4). The results were used to generate an improved User Interface (UI) being proposed (figure 5) that was tested by focus groups of independent university students.

The results indicated that some of the users asked why it was impossible to move the clothes left and right when they could move them vertically. They requested the option to select other clothes directly from the system, without having to return to the main page (which in the LazyLazy system requires going back to the computer and using the mouse. The system should be able to remember the size of an adjusted piece of clothing, as well as recognize the size of the user in general. Furthermore, the users would like to be able to view accessories (e.g. watches, hats, glasses) together with the clothes. The UI should allso offer the possibility to view the back of a piece of apparal (especially since some have patterns or pictures on the back). Some users also asked for the possibility of using voice control to interact with the system.

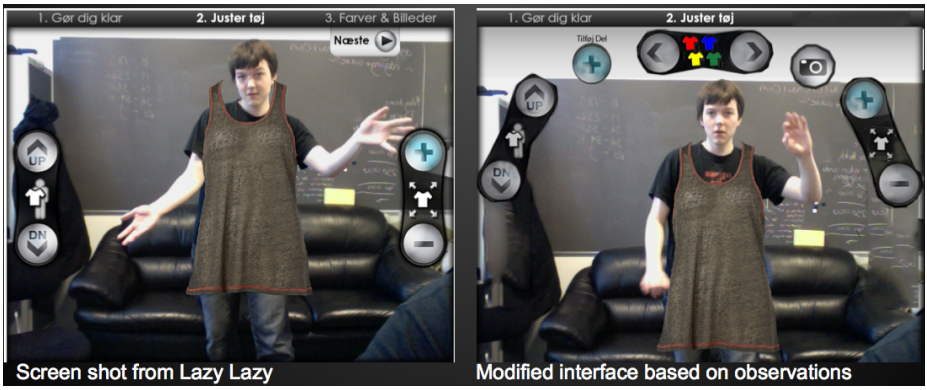


Fig. 5. Existing interface (left) and proposed interface (right) as tested by focus groups

Figure 5 shows a proposed change to the LazyLazy UI with button relocation and two steps combined into one. Three focus group interviews were conducted with 13 participants, approximately 30 minutes per group. Feedback was that the existing interface was considered as less overwhelming for new users; incrementaly takes the users through step-by-step; and is very guiding. The proposed interface was considered as allowing modifications of all aspects in one step; buttons were easier to reach; and looked nicer. Both were judged to need improved and larger text. The preferred design was 3/10 in favor of the proposed change of UI design.

4 Conclusions

Online shopping means not having to deal with crowds or pushy/inattentive sales personnel. It also saves money on travel expenses (car, petrol, parking…) and ease of getting the goods through your door via direct postage/freight delivery that means no 'mountain mule' shopping (or public transport) with numerous bags encumbering the person. Thus, potentials for an optimal system are huge, however, many constraints

limit such optimization, mostly feel and fit aspects. The goal of this project is to advance the state of the art via an optimal dynamic system that would reduce purchased apparel returns.

Design options in such work include for the user to control a matched avatar through body gesture. However, this has not been implemented in the VDR project where the focus is on a mirrored representation with a superimposed overlay using camera recognition and mapping.

A challenge, known and strengthened according to the PD studies to date, is the feel and fit aspects. In line with this are the technical challenges involved in cloth dynamics (mirrored) simulation where the virtual matches the physical such that interactive rate correspondences between the 2D and 3D views and the simulation addresses geometric nonlinearity and frictional contact while remaining stable even under rapid user input. The Sensitive Couture tool [6] claims to address such issues by combining techniques including (i) fast prediction of 3D forms from cached shapes using sensitivity analysis and generalized moving least squares, (ii) fast invisible re-meshing using positive mean value coordinates to accommodate arbitrary revisions of the pattern boundary, and (iii) stable and accurate cloth modeling using an isometric bending model, a modified St. Venant- Kirchhoff membrane element, and progressive refinement.

PD outcomes also point to size of monitoring screen being problematic with a need to balance distance to view full body comfortably and ease of interaction with UI buttons. To help eliminate issues such as illustrated in figures 2 & 3, future research is posited to explore using large screen portrait orientation (vs traditional landscape orientation) monitors that enable real-time 3D auto stereoscopy[12].

Wheelchair bound interviewees made clear their need for improved means to purchase clothes without having to leave their homes. The perceived ease of use (PEOU) and perceived usefulness (PU) toward enhancement and reduced effort of an accessible online apparel purchasing system were key factors when wheelchair-bound consumers were questioned. This highlighted how a market opportunity is clearly evident for people in wheelchairs to augment accessibility when purchasing for self or partners, family, or friends.

Conflicts of Interest Statement. The authors have no conflict of interest to disclose.

Acknowledgements. The Virtual Dressing Room project is financed by the Danish National Advanced Technology Foundation, Denmark.
The researchers want to acknowledge the host companies in UK (section 3.1); Lazy-Lazy; and Bruuns Galleri, Aarhus.

References

1. Brandt, E.: Designing Exploratory Design Games. A Framework for Participation in Participatory Design? In: Proceedings from Participatory Design Conference 2006, pp. 57–66. ACM (2006)

[12] http://www.tridelity.com

2. Buur, J., Bødker, S.: From Usability Lab to "Design Collaboratorium": Reframing Usability Practice. In: Proceedings of DISoo: Designing Interactive Systems: Processes, Practices, Methods, & Techniques, pp. 297–307 (2000)
3. Buur, J., Jensen, V.M., Djajadiningrat, T.: Hands-only Scenarios and Video Action Walls; Novel Methods for Tangible User Interaction Design. In: Proceedings of DIS 2004: Designing Interactive Systems: Processes, Practices, Methods & Techniques, pp. 185–192 (2004)
4. Muller, J.M.: Participatory Design: The Third Space in HCI. In: Proceedings of the Participatory Design Conference 2000, pp. 147–155 (2000)
5. Seifert, M.C., et al.: Demystification of Cognitive Insight: Opportunistic Assimilation and the Prepared-Mind Persopective. In: Sternberg, R.J., Davidson, J.E. (eds.) The Nature of Insight. The MIT Press, Cambridge (1995)
6. Umetani, N., Kaufman, M.D., Igarashi, T., Grinspun, E.: Sensitive couture for interactive garment modeling and editing. ACM Trans.Graph 30(4), Article 90 (2011)

Digital Menu Boards as Influencer for Healthy Eating

Anicia Peters[1] and Brian Mennecke[2]

Iowa State University
[1] 1620 Howe Hall,
[2] 3313 Gerdin Building,
Ames IA 50011, USA
{anpeters,mennecke}@iastate.edu

Abstract. This paper reports on research designed to examine the role of embedded videos in digital menu boards and it's influence on consumer decision-making. The study's results showed that the embedded videos caught the attention of decision makers and healthy food images have an effect on choice; however, these variables do not interact. The results for the main effects show that technology can aid decision-making during the information search and purchase/choice stages.

Keywords: Digital signage, digital menu boards, decision-making, healthy eating, information overload, video boards.

1 Introduction

A common challenge for advertisers is to capture the attention of consumers. Nevertheless, this challenge is growing more daunting as consumers are presented with an increasing number of advertisements from various sources and via numerous media. For example, ads are presented to prospective customers on signs in public spaces, on mobile phones, and on computers and a common result is that the intended recipient of the message copes with this deluge by ignoring most of what they are exposed to. This coping strategy is called "display blindness" [14] and an unfortunate result is that the consumer may throw aware the good with the bad by ignoring information that might be useful in helping him make better decisions [7].

This problem is most pronounced in public spaces and at retail venues where signs are used both for advertisement and also for information dissemination. One solution that has arisen in the last few decades is to use digital signage, which because of lower costs and increasing capabilities, offers advertisers the ability to present information in more dynamic and vivid ways with the goal of attracting the attention of consumers [3]. Digital signage used for advertising using consists of networked flat panel display screens capable of displaying multimedia and high-speed Internet feeds [3].

For example, a number of fast-food retailers such as Wendy's and McDonalds have deployed dynamic digital menu boards [18] as is seen in figure 1 to attract customer attention and promote their products and promotions. These displays often prove to be quite effective. For example, Wendy's reported an upsurge in sales due to

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 428–437, 2013.

the use of menu boards at the point of sale [18]. Of course, while fast-food venders may have initially been motivated to use these technologies to increase sales or push high-margin products, with the increasing interest in healthy eating and forthcoming requirements for presenting nutritional information [15] it is likely that digital menus will also provide potential benefits beyond increasing the bottom line. For example, schools have begun to use digital signage in their lunchrooms and dining centers [8, 16] in an effort to promote healthy eating. Given this, we suggest that these technologies could be useful in enticing consumers to make healthier eating choices in point of sale settings. Yet, little has been done to examine whether and how signage influences decision making at the point of sale; therefore, our focus is on examining whether these menu boards can be used to influence healthy choices.

Fig. 1. A dynamic digital menu board at Wendy's

2 Research Question

Our research question focuses on whether the video in dynamic digital menu boards can influence consumer decision in such a way that consumers will make healthier food choices when confronted with video food ads of healthy food choices.

3 Literature Review

3.1 Information Overload

One of the characteristics of the modern world is the fact that technology enables businesses, governments, and other people to "reach out and touch" us as we carry out our daily activities. In some cases, we "subscribe" to services that "touch" us with news feeds, promotions, or personal messaging, but as often as not we are also presented with unsolicited messages that arrive via smartphones or through displays we encounter in our environment. The cumulative effect of this information deluge is that recipients of this messaging are often overloaded, which makes filtering through the noise to hear the "signal" buried within the noise increasingly difficult and this can also negatively affect decision making [1]. In other words, because consumer's have a limited working memory capacity and are often pressed for time, they are likely to experience "information overload" [1, 13], which occurs "…when too much

information affects a person and the person is unable to recognize, understand or handle this amount of information…" [7].

Consumers tend to develop coping strategies in reaction to excess stimuli. For example, some of the coping strategies that are suggested in the literature are that consumers process information less in-depth, they ignore less relevant information, they delay or abandon reaching a decision, and they make less optimal decisions [7, 11]. Other studies have found that consumers often learn to ignore stimuli such as signage and displays [14]. An important question is, what can break through to the consumer? In the context of digital signage displays in public venues, one solution that many advertisers are using is to imbed video, sound, movement, and other vivid elements into displays with the goal of catching the involuntary selective attention of consumers [10, 14]. Of course, an important empirical question yet to be addressed is what factors in a display will be effective in overcoming display blindness and garnering the attention of customers? Our research addresses this question in part by examining one dimension of vividness, the movement of objects on the display.

3.2 Consumer Decision-Making

To examine the question of the role of vividness in overcoming display blindness, we examine the role of movement in a context similar to the point of sale context of fast-food restaurants. We first encountered digital signage used for purveying information about food at such an establishment, and the venue is ideal for examining whether and how the display characteristics influence attitudes and behaviors. Of course, the decision about whether and what to eat is complex and will be influenced by more than merely a sign in a restaurant. For example, consumers go through multiple stages when deciding whether and what to purchase. One of the standard models of consumer decision making is the Engel, Blackwell and Miniard (EBM) model, which offers seven stages consisting of *need recognition*, *search*, *pre-purchase alternative evaluation*, *purchase*, *consumption*, *post consumption evaluation* and *divestment* [6]. The EBM model distinguishes between individual differences, environmental influences and psychological processes that influence consumer decision-making through information processing, learning and attitude and behavior change.

While the EBM model is useful in a number of contexts, not all consumers go through all the decision-making stages and decision rules. Specifically, stages are used or skipped depending on the degree of complexity of the choice and the consumer's involvement in the process during each stage [6, 9, 19]. Additionally, a number of other factors influence which stages are processed, how they are processed, and how the processing influences outcomes. For example, problem size is important and is determined based on the number of options to be evaluated within the available time. Furthermore, time pressure causes consumers to adapt their decision strategies according to the amount of time available for a decision [2]. Furthermore, information format plays a significant role in decision making as shown by studies suggesting that complex information structures on display causes increased cognitive load, thus influencing the way that information is processed during each stage [1, 2].

The format of display features in particular can have an important influence on how information is processed. Many menu boards are structured with items presented in a list and therefore the consumer can view all of the information simultaneously, which should have a positive impact on interpretation and decision-making [1].

Nevertheless, many menus are still cluttered with too much information, which makes interpretation difficult [12]. One way that has been presumed to draw attention to relevant content is to increase the vividness of contents on the display [10, 12]. Nevertheless, while movement, color, and similar vivid elements have been shown to be influential in contexts such as web site ads, little has been done to identify how these elements influence decision making in the context of menu boards.

3.3 Healthy Eating

Fast food restaurants have recently come under fire because of the association between fast food consumption and obesity [4]. Nevertheless, many of the items sold at fast food restaurants are quite healthy and if venders could encourage increased consumption of these items not only would the health of their patrons improve, but so too would the vender's bottom line given that healthy items often have higher margins.

Digital menu displays offer potential to encourage healthy eating. For example, digital menus can be used to display nutritional information along with information about the product. In fact, restaurants in the US are now required to display nutritional information on their menus as required by the 2010 Health Care Reform Act [15]. Nevertheless, displaying more information on an already overcrowded menu board will undoubtedly increase the information load of consumers and might result in exactly the opposite outcome of what is desired; that is, consumers might become blind to this nutritional information. Understanding the role of information content and presentation format will be important in helping venders present nutritional content to consumers and potential in increasing healthy eating behaviors.

4 Framework and Research Model

We considered the basic five stages in consumer decision-making as presented by [5] as: a) need recognition, b) search, c) alternatives evaluation, d) choice/purchase, and e) outcomes. We also considered other factors such as the type of consumer, situational factors, relevance of the ad, personal preferences, motivation, information overload, and display layout conditions. We examined hypotheses associated only with the *information search stage,* the *alternatives evaluation stage*, and the *purchase/choice stage.*

To derive the hypotheses, it is useful to present them in the context of the five basic consumer decision-making stages [5] as discussed above. We offer the research model in Figure 2 to summarize our hypotheses as we discuss in detail below.

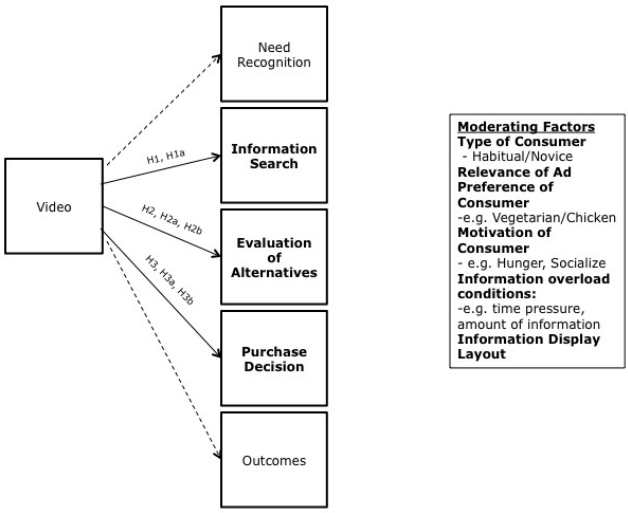


Fig. 2. Research model

Need Recognition Stage: We will not consider the need recognition stage in our research because the scenario in which our subject is placed in our study is at the Point-of-Purchase (POP) (i.e., the subject is presented with the video menu board immediately before the decision is made). Future studies could examine the role that post-consumption evaluation plays in the need recognition stage.

Search Stage: This stage is characterized by an external and internal search for information. The internal search source is prior knowledge, whereas the external search will be primarily limited to the video menu as this is typically installed at the Point-of-Purchase (POP). The nature of the information search is influenced by goals, time, financial resources, and sensory information. Information processing (i.e. exposure, attention, comprehension, acceptance, and retention) is a very important part of the search stage [6]. We expect that subjects will pay *attention* to the video ads because it serves to declutter the displays and attract selective attention through its vividness, colors and moving imagery. Because video ads are in line of sight at the POP, consumers might be unable to screen out the video completely and by holding the consumer's attention, *comprehension* and *acceptance* will be greater. Memorable video ads will also more likely be *retained* in memory and serve as a future internal memory trigger when consumers recall information from previous exposure to the ads. Our hypotheses for this stage are thus:

H1: A video in a dynamic digital menu board will draw more attention than an image in a static menu board.

H1a: Moving imagery via video in a dynamic digital menu board will be more salient than an image in a static menu board.

Evaluation of alternatives: According to previous studies [1, 2], variables such as problem size, time pressure, attribute correlation, completeness of information, information format, and comparable versus noncomparable choice all influence the evaluation of alternatives. Because the video ad will be evaluated alongside static text/graphics and internal memory, we propose the following hypotheses:

H2: The video in a dynamic digital menu board will lead to the evaluation of fewer alternatives than an image in a static menu board.

H2a: The video in a dynamic digital menu board will reduce the consumer's perception of decision complexity more than an image in a static menu board.

H2b: The video in a dynamic digital menu board will lead to higher levels of involvement than an image in a static menu board.

Choice/Purchase: A consumer may employ several decision-strategies or combinations thereof depending on the situation or depending on the choice complexity [1, 2]. Although choice complexity is influenced by a number of factors such as the number of alternatives and attributes, the processing difficulty, the uncertainty, the characteristics of the person, and the social context (e.g., family decision-making versus individual choices) [1], the video ads will play a significant role in influencing the final decision. Therefore, the following hypotheses are proposed.

H3: The video in a dynamic digital menu board will become a reference heuristic for decisions more than an image in a static menu board.

H3a: The availability bias of the video in a dynamic digital menu board will increase the likelihood of the featured item being purchased more than an image in a static menu board.

H3b: The persuasiveness of the video in a dynamic digital menu board will increase the likelihood of the featured item being purchased more than an image in a static menu board.

Outcomes/Post-purchase alternative evaluation: When a consumer evaluates the product after consumption, this can influence future decision-making by associating the video ad with a positive experience. The video will then, in turn, serve as positive recommender for future purchase decisions. We did not test this stage.

5 Methods

We conducted a 2 x 2 factorial study with independent variables of health condition (healthy, unhealthy) and visuals (video, static). We simulated a menu board consisting of still images and video clips of two unhealthy items, two healthy items, and eight additional food items displayed in plain text.

A total of 134 undergraduate students participated in the research, of which 40% were male. The video clips consisted of moving images of selected food items (i.e., they were randomly shown either less healthful items such as chips and cookies or more healthful items such as peanuts or raisins). The food items were selected and classified by having a panel of subjects rate the items on appeal and perceived healthfulness. Movement in the videos was created using the Ken Burns effect, which involves panning across the object to show details and to create a sense of movement.

Subjects were given 20 seconds to view the menu choices (figure 3) with the video clips and make a selection of a food product to eat. Following their selection, subjects were asked to complete a survey.

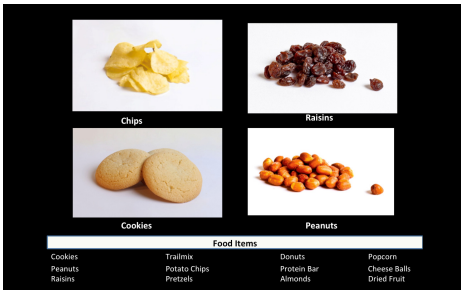


Fig. 3. Menu board simulation during study

6 Results

ANOVAs were conducted on composite structures to test our hypotheses. Covariates measuring habitual choices and time pressure were included in the analysis.

Information Search Stage (H1 and Ha): Main effects were found for both visuals (p=.000) and healthy choices (p=.007), but no support for an interaction between visual and healthy choices were found. The habitual choice covariate had an influence on the effect size.

Hypothesis 1 (the video will catch the observer's attention more than static images) showed main effects for visuals (p=.008). Since the effect size for healthy condition was low (R^2 =.144), the power analysis (.491) showed the healthy condition would have shown a main effect if there were more subjects. The habitual choice covariate had an influence on the main effect. No support for interaction was found.

H1a (motion of video attracts attention more than static images) showed main effects for visuals (p=.000) and healthy choices (p=.005). The habitual choice covariate had an influence on the effect size. However, there was no support for the interaction between visuals and healthy choices.

Alternatives Evaluation Stage (H2, 2a, 2b): For this stage, no main effects or an interaction effect was encountered. The covariate habitual choice increased the effect slightly to R^2=.071. H2 was not supported. No main effects were determined for either visuals (p=.341) or healthy choices (p=.418). No interaction effect was

observed. H2a (reduction of perception of decision complexity) showed no main effects for either visuals (p=.581) or healthy choices (P=.940) and there was also no interaction effect. H2b (high involvement level) showed a main effect for only visuals (p=.007) and none for healthy choices (p=956). No interaction effects were demonstrated.

Purchase/Choice Stage: This stage showed that consumer choice can be influenced by video (main effect p=.020) and by healthy options (main effect p=.017). Nevertheless, no interaction effects were found. None of the covariates had a significant effect on the group hypothesis for the purchase stage. H3 (reference heuristic) showed no main effects for visuals (p=.118) or healthy choice (p=.057). No interaction effects were evident. H3a (availability bias) was not supported at all with no main effects (visuals: p=.216 and healthy choice: p=.133) and no interaction effects. Nevertheless, two covariates, "time was too short" and "habitual choice," influenced the effect size (R^2=.173). A power analysis revealed a larger sample size might increase the significance of the main effect healthy choice. H3b (persuasiveness of vividness effect) showed main effects for visuals (p=.007) and healthy choices (p=.014), but no interaction effects was demonstrated.

Full model ANOVA: The full model showed main effects for both visuals (p=.003) and healthy choices (p=.014), but again no interaction effect was observed. Habitual choices showed a significant influence on the effect size (R^2=.167).

7 Discussion

For the overall model, our ANOVA results showed main effects for visuals (p=.003) and healthy choices (p=.014) with a significant covariate for habitual choices, but failed to show an interaction effect. The smaller samples size and the low effect sizes resulted in no significant interaction effect; therefore, a larger sample size might show results for the interaction.

Another explanation for the lack of interaction is that subjects did not consider the health aspects of the snacks, as their choices did not correspond with their self-reported health-consciousness. Nevertheless, their general perceptions of whether they chose a healthy item agreed with their self-report health rating of the item.

Another factor influencing the results relates to the fact that subjects mostly made habitual choices, i.e. they selected food items that they usually eat (64%) and usually those that they usually like to eat (83%). As much as this indicates a familiarity bias in the study, it also indicates a habitual decision-making style, which means that they did not go through pre-defined steps in the decision process [9, 19]. By using this decision-making style, subjects seemed to have used prior knowledge to guide their decisions and blocked new information or unfamiliar products [9, 17]. This situation was further fueled by the fact that subjects' involvement levels were very low as they participated in the study primarily because they would earn research credits. The habitual decision-making style is also evident in the alternatives evaluation stage hypotheses where no main effects were found for either visuals or healthy eating.

The good news is that the video and health aspects of the food had significant effects during the information search stage and the purchase/choice stage, which indicates that this technology can be used to influence decision-making.

8 Conclusion

Although we found no evidence for an interaction effect between video use in digital menu boards and healthy eating, our results demonstrate main effects for video and healthy eating choices. We thus suggest that the videos can be used to aid decision-making and influence healthy food choices.

References

1. Bettman, J., Johnson, E., Payne, J.: Consumer decision making. In: Robertson, T.S., Kassarjian, H.H. (eds.) Handbook of Consumer Behavior, pp. 50–84. Prentice Hall Inc., Englewood Cliffs (1991)
2. Bettman, J., Luce, M., Payne, J.: Constructive consumer choice processes. Journal of Consumer Research 25, 187–217 (1998)
3. Burke, R.: Behavioral effects of digital signage. Journal of Advertising Research 49(2), 180–186 (2009)
4. Drewnowski, A., Darmon, N.: The economics of obesity: dietary energy density and energy cost. American Journal of Clinical Nutrition 82(1), 2655–2735 (2005)
5. Engel, J., Blackwell, R., Kollat, D.: Consumer Behavior, 3rd edn. Dryden, Hinsdale (1978)
6. Engel, J., Blackwell, R., Miniard, P.: Consumer Behavior. The Dryden Press, Harcourt Brace College Publishers (1995)
7. Eppler, M., Mengis, J.: The concept of information overload: A review of literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines. The Information Society, An International Journal 20(5), 1–20 (2004)
8. Glasgow City Council serves up healthy eating message to city's children using ONELAN digital signage, `http://www.mxdigitalsystems.com/documents/Onelan_Glasgow_Healthy_Eat_case_study.pdf`
9. Hoyer, W.: An examination of consumer decision making for a common repeat purchase product. Journal of Consumer Research 11, 822–829 (1984)
10. Huang, E.M., Koster, A., Borchers, J.: Overcoming assumptions and uncovering practices: When does the public really look at public displays? In: Indulska, J., Patterson, D.J., Rodden, T., Ott, M. (eds.) PERVASIVE 2008. LNCS, vol. 5013, pp. 228–243. Springer, Heidelberg (2008)
11. Jacoby, J., Speller, D., Kohn, C.: brand choice behavior as a function of information load. Journal of Marketing Research 11, 63–69 (1974)
12. Jansson, C., Marlow, N., Bristow, M.: The influence of color on visual search times in cluttered environments. Journal of Marketing Communications 10, 183–193 (2004)
13. Malhotra, N.: Reflections on the information overload paradigm in consumer decision making. Journal of Consumer Research 10, 436–440 (1984)

14. Müller, J., Wilmsmann, D., Exeler, J., Buzeck, M., Schmidt, A., Jay, T., Krüger, A.: Display blindness: The effect of expectations on attention towards digital signage. In: Tokuda, H., Beigl, M., Friday, A., Brush, A.J.B., Tobe, Y. (eds.) PERVASIVE 2009. LNCS, vol. 5538, pp. 1–8. Springer, Heidelberg (2009)
15. Rosenbloom, S.: Calorie Data to be posted at most chains. The New York Times (March 23, 2010), http://www.nytimes.com/2010/03/24/business/24menu.html
16. School District of Philadelphia Uses Digital Signage to Promote Get Healthy Philly, http://www.digitalsignageconnection.com/3School-District-of-Philadelphia-Uses-Digital-Signage-to-Promote-Get-Healthy-Philly86
17. Stijn, M., Van Osselaer, J., Alba, J.: Consumer learning and brand equity. Journal of Consumer Research 27, 1–16 (2000)
18. The Buzz: Installation spotlight: POP appeal, http://svconline.com/digitalsignage/features/digital_signage_wendys_0603/
19. Wright, P.: Consumer choice strategies: Simplifying vs Optimizing. Journal of Marketing Research XII, 60–67 (1975)

Reverse Business Innovations – The Impact of ERP Software Upgrades on Organizations

Stefanie Rauff[1] and Andreas Hufgard[2]

[1] University of Wuerzburg
srauff@wiinf.uni-wuerzburg.de
[2] University of applied sciences Aschaffenburg
hufgard@ibis-thome.de

Abstract. Vendors of enterprise software continue to evolve their products and offer user organizations the latest functionality in new releases and upgrades. But this means companies are faced with the challenge of selecting features that add value, and trying to assess how implementation would impact them. Not every new function impacts all employees in the same way. Depending on the specific organizational structure, implementation will affect a varying number of users. And depending on a company's system landscape, implementation of new functions might require a range of IT activities and affect many processes.

The authors have developed a concept that uses an organization's system usage data to deliver decision support. Various aspects of the upgrade can be assessed with the aid of software tools (analyses and evaluations of new functionality) to determine relevancy, benefits, profitability, strategic goals and impact. In this article we will present a portion of this concept: A model that can assess how new functionality would impact a business, if implemented.

Keywords: SAP ERP, upgrade, new release functions, decision support, usage analysis.

1 Research Topic

An enterprise system (ES) is defined as a standardized software package whose aim is to integrate the whole of an organization's data [1]. ES development can be described in two waves [2]: The first consists of procurement, configuration and implementation. The second, also commonly referred to as the post-implementation phase, focuses on achieving continuous improvement and driving value. Questions central to this phase include how to gain greater benefit from the investment, how to manage and enhance the ES and how the ES will impact the organization.

1.1 Current Research in the Post-implementation Phase

The first wave has been widely researched [3-8]. A number of surveys and case studies have been done on the implementation wave, with respect to strategic options,

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 438–445, 2013.

failure avoidance, identification of alignment issues and improved implementation methodologies [9].

Recently, research has moved beyond implementation to the second wave [2]. Research topics include: the impact of ES implementation on financial performance [10] and [11]; standardization and trust relations [12]; legitimization of certain practices in the organization [13]; continuous changes in business processes [14]; potential impact of enterprise system implementation and its effects on the organization [15]; user acceptance and expectations [16]; the impact of ES implementation on job satisfaction [17] and on the organization [18], and how employees' social network structures affect post-implementation success of enterprise systems [19]. In short, they deal mainly with various aspects of implementation and how these affect organizations.

This paper also concerns the second wave, but is not devoted to the impacts of implementation. Its focus is on the variety of new functions regularly provided by enterprise software vendors in upgrades.

In this paper we will examine specific types of innovations that are characterized as follows:

- They are functional innovations
- and are included in an upgrade for enterprise software.

To set this specific type of innovation apart, we invented the term "upgrade innovation."

In this paper, the research question is: How can we help organizations evaluate the impact of upgrade innovations?

1.2 Research Methodology

The model was developed in the following steps:

1. The current situation and problem were analyzed (section 1)
2. A prototype was created (sections 2 and 3)
3. The model was validated (section 4)

An analysis of relevant literature revealed that next to no theory exists on upgrades. The subject of untapped potential in business upgrades has been neglected entirely.

We created a prototype and identified assessment criteria (relevancy, cost-benefit analysis, impact and strategic goals) according to which upgrade innovations were to be assessed. This paper will present the issue of impact assessment.

2 Description of Method

We identified a lack of research on evaluating the relevancy of new functions offered in upgrades, and on system-supported impact assessments prior to an upgrade.

A vast number of organizations that use enterprise systems are confronted with the same problem. This is why it makes sense to build an analysis tool for this purpose

rather than let each company review all innovations by itself. The authors of this paper are currently collaborating with SAP to create the technology and content necessary for this tool.

2.1 Reverse Business Innovations

Our goal was to develop a model that would help organizations assess the impact of upgrade innovations relevant for them. To this end, we looked for criteria that would enable objective inference of an upgrade innovation's impact. In other words, we posed the question: Under which conditions would this functionality have an effect on the company? We developed an approach we dubbed "Reverse Business Innovations."

Figure 1 shows how the system-based usage analysis and the rule-based impact assessment are interrelated.

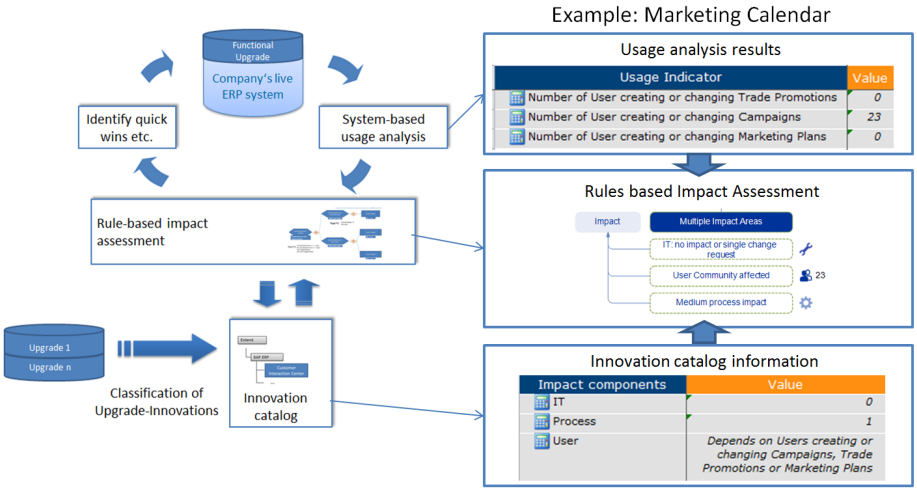


Fig. 1. RBI model

In preparation for the analysis, we examined and classified the upgrade innovations to gain insight into the new functionality offered. For each innovation, we identified criteria that enabled us to infer the impact of the innovation, i.e. what technical requirements an organization would have to fulfill to enable us to conclude how the innovation would impact it. We then filed away this information in the "innovation catalog," a reference for determining, for example, which indicators need to be mined from an organization's system usage data.

When the analysis is conducted for an organization, the usage indicators are extracted from the company's live ERP system. They help determine impact based on the innovation catalog. The results enable us to assess the impact an upgrade innovation has on an organization and so, immediately identify quick wins.

2.2 Impact Assessment

In their maintenance model, NG et al. indicate that a gap analysis must be conducted between the new upgrade and the existing software version in order to identify the impact on an organization, and finally, to create a budget, a schedule and a plan for employees [6].

Organizations are often short of resources [20], so an impact assessment can help them identify upgrade innovations that require only very few resources.

We defined three distinct aspects of the term "impact":

- Effect on users: The number of employees who will use the new functionality,
- Effect on IT: This can include anything from a simple change request to new hardware or software,
- Effect on processes: This considers the range of processes for the organization

Employees

An upgrade always entails change, so it is essential to incorporate routine change management activities, such as risk management assessments, training, and knowledge transfer [21] and [3]. In a qualitative study, Zhao found a correlation between engaging employees in the upgrade and offering them training, and an upgrade's success [22]. So we can conclude that it is essential to consider the impact on the employees of an organization when contemplating an upgrade.

This means it is important to conduct an assessment of the number of people impacted by the upgrade innovation. We defined three categories for impact on users:

- Few users affected
- (Individual) user communities affected
- Many users affected

IT

According to Brose, the extent of the risk posed by an innovation can be determined by the following two factors [23]:

1. The degree of technical difficulty, because it is a measure of the ability to solve a problem that occurs, especially during the required development time.
2. The technical complexity, because it is a measure of the diversity and number of individual activities that must be coordinated.

Since upgrade innovations are not products or services the company itself has developed, but software tested and supplied by a vendor of a business-related information system, we can disregard the first factor when assessing upgrade innovations. But the second factor involves the number of technical measures that must be implemented and how complex these are, and so, must be included in the impact assessment.

We defined three categories for technical measures:

1. A single change request
2. The IT resources required (developers, project managers) for a change project
3. New hardware or software required

Normally this value is set by the manufacturer of enterprise software, because it is not dependent on the specifics of any one organization.

Processes
Companies are predominantly organized by processes. And these might range across several user departments or even across companies. Organization by processes also enables problems to be solved across functions [24].

Therefore, when examining impact, it is important to look at the effects of upgrade innovations on the specific process in question. We defined three categories for impact on processes:

1. Affecting isolated processes
2. Affecting several processes
3. Affecting a large number of processes

Normally this value is set by the manufacturer of enterprise software, because it is not dependent on the specifics of any one organization.

3 Validation

We will conclude our paper with a validation of the model based on a concrete upgrade innovation from SAP CRM.

We created a prototype in a rule-based expert system specifically for the new upgrade functionality in SAP CRM added between Release 5.0 and 7.0 EHP 2. We then validated it on the system of a manufacturing company that operates worldwide. The CRM system examined is strategically crucial for this user organization. Approximately 4,000 users from Europe, Asia and North America actively work within it and it links three ERP systems.

The company currently uses Release SAP CRM 7.0 EHP 1. They last upgraded because of system consolidation with a subsidiary whose CRM system used EHP 1.

The pilot organization agreed to the analysis because decision-makers thought some of the necessary SAP CRM areas may not have been activated. In addition, they were interested in obtaining new functionality available in Release 7.0 EHP 2 without much added expense.

Usage analysis results showed that during the analysis period, 232 campaigns were used by 23 employees. Leads, trade promotions and marketing plans were not used at all. Trade promotions are used to depict marketing activities in collaboration with partners. Marketing plans track the activities in Marketing at a higher strategic level and for a longer period than trade promotions and campaigns.

The marketing calendar is an upgrade innovation that offers users a graphic overview of all marketing projects according to their duration. Projects can also be created directly on the graphic.

Under **impact** it shows that for IT, the marketing calendar has very little impact (see Figure 2), because only one function (i.e. the business function) has to be

activated. The users affected are the employees that create the marketing projects (campaigns, trade promotions and marketing plans). They total 23. The marketing calendar's impact on processes has been rated medium, because processes concerning trade promotions and marketing plans might also be affected.

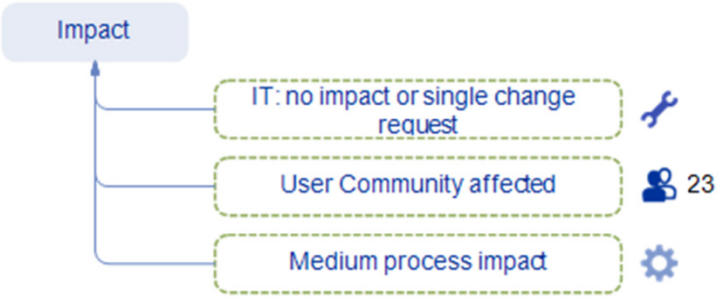


Fig. 2. Impact assessment for the marketing calendar

Overall, impact assessment reveals that several impact areas are affected: 23 users and a medium impact on processes.

4 Discussion of Results

The newly developed model delivers organization-wide decision support that determines the impact of an upgrade innovation – automatically and according to rules.

In sum, the new model offers a number of benefits to user organizations, and also to manufacturers of enterprise software.

User organizations benefit most by the use of automated assessment when comparing their as-is situation to their (potential) situation following an upgrade:

- They save time otherwise spent gathering and assessing information for an upgrade.
- They require less external consulting and so, reduce costs.
- Consultants' advice/recommendations are replaced by computer-based activities and so, reducing the number of errors made.
- Recommendations are easier to understand because detailed data and explanations are included in the results.
- The computer-based analysis is more thorough than an analysis conducted by human consultants. It also recommends upgrade innovations from earlier releases that may have become relevant under current company conditions.
- The tool can be used on local systems and so, is always available. It can also be employed in several user organizations simultaneously.
- The expert system delivers much faster results than human experts can.
- The tool contributes to long-term transparency, because it is very simple to document the results in detail.

The consequence of all above-mentioned points is that user organizations will begin upgrading sooner to higher releases and the upgrade innovations will be put to use and their potential tapped sooner. In some cases, the upgrade innovation can replace a customer-specific add-on with a standard function. All of these consequences would result in far lower costs for maintaining the enterprise software.

What is more, vendors of enterprise software would also benefit from the availability of this kind of tool:

- User organizations would no longer rely as much on the expensive maintenance of old releases if they had an incentive to upgrade.
- User organizations would require less consulting for the same or better service, and so, improve their cost structure (copying software costs nothing).
- Software vendors would enjoy improved customer loyalty because organizations would be able to benefit more from their enterprise software (and changing software would be more expensive and time-consuming).

In light of the prevalence of enterprise software, the wealth of practical applications for the model presented in this paper is clear. This tool could greatly reduce costs for both user organizations and manufacturers, and could improve the integration of upgrade innovations in organizations.

This same methodology can also be applied toward new products not included in an upgrade, but that require additional licensing costs.

In conclusion, it is essential to address the issue of how the analysis results can accelerate and prepare user organizations for the impending implementation and application of any given upgrade innovation.

References

1. Davenport, T.H.: Putting the enterprise into the enterprise system. Harvard Business Review, 121–131 (July/August 1998)
2. Kræmmergaard, P., Koch, C.: Managing ERP after going-live. In: Proceedings of the 9th International Conference European Operations Management Association, Copenhagen (2002)
3. Esteves, J., Pastor, J.: Towards the unification of critical success factors for ERP implementations. In: Proceedings of the 10th Annual BIT Conference (2000)
4. Dong, L., Neufeld, D., Higgins, C.: The Iceberg on the Sea: What do you see. In: Proceedings of the 8th Americas Conference on Information Systems, Paper 124, pp. 857–864 (2002)
5. Jacobs, F., Bendoly, E.: Enterprise Resource Planning: Development and Directions for Operations Management Research. European Journal of Operational Research 2(146), 233–240 (2003)
6. Ng, C., Gable, G., Chan, T.: An ERP Maintenance Model. In: Proceedings of the 36th Hawaii International Conference on System Sciences (2003)
7. Moon, Y.: Enterprise Resource Planning (ERP): a review of literature. International Journal of Management and Enterprise Development 4(3) (2007)

8. Maedche, A., Mueller, B.: Enterprise Systems - A Research Agenda (2012), http://ssrn.com/abstract=2033145, http://dx.doi.org/10.2139/ssrn.2033145
9. Al-Mashari, M.: Enterprise resource planning (ERP) systems: a research agenda. Industrial Management & Data Systems 1(103), 22–27 (2003)
10. Poston, R., Grabski, S.: Financial impacts of enterprise resource planning implementations. International Journal of Accounting Information Systems 2, 271–294 (2001)
11. Hunton, J., Lippincott, B., Reck, J.: Enterprise resource planning systems: comparing firm performance of adopters and nonadopters. International Journal of Accounting Information Systems 4, 165–184 (2003)
12. Locke, J., Lowe, A.: Standardisation and trust relations in ERP implementation: reflections from a case investigation. In: 28th Annual Congress of the European Accounting Association, Goeteborg, Sweden, May 18-20 (2005)
13. Dillard, J., Ruchala, L., Yuthas, K.: Enterprise resource planning systems: a physical manifestation of administrative evil. International Journal of Accounting Information Systems 6, 107–127 (2005)
14. Davenport, T.H., Brooks, J.D.: Enterprise systems and the supply chain. Journal of Enterprise Information Management, 17(1), 8–19 (2004)
15. Rikhardsson, P.: Kræmmergaard: Identifying the impacts of enterprise system implementation and use. International Journal of Accounting Information Systems 7(1), 36–49 (2006)
16. Liang, H., Saraf, N., Hu, Q., Yajiong, X.: Assimilation of Enterprise Systems: The Effect of institutional pressures and the mediating role of top management. MIS Quarterly 31(1), 59–87 (2007)
17. Morris, M.G., Venkatesh, V.: Job Characteristics and Job Satisfaction: Understanding the role of ERP Implementation. MIS Quarterly 34(1), 143–161 (2010)
18. Volkoff, O., Strong, D.M., Elmes, M.B.: Understanding enterprise systems ‐ enabled integration. European Journal of Information Systems 14(2), 110–120 (2005)
19. Sasidharan, S., Santhanam, R., Brass, D.J.: The Effects of Social Network Structure on Enterprise Systems Success: A Longitudinal Multilevel Analysis. Information Systems Research, 1–21 (2011)
20. Cooper, R.: Top oder Flop in der Produktentwicklung. 2. Aufl. Wiley-VCH Verlag, Weinheim (2010)
21. Hernández, J., et al.: The SAP R/3 handbook. 3. Aufl. McGraw-Hill, New York (2006)
22. Zhao, F.: ERP Upgrade Report (2007)
23. Brose, P.: Planung, Bewertung und Kontrolle technologischer Innovationen. Erich Schmidt Verlag, Berlin (1982)
24. Krüger, W.: Grundfragen der Bildung und Integration organisatorischer Subsysteme. In: Franz, X., et al. (eds.) Allgemeine Betriebswirtschaftslehre Bd.2: Führung. 8. neubearbeitete und erweiterte Aufl., Lucius & Lucius, Stuttgart, pp. 127–157 (2001)

Repository-Based Implementation of Information Pyramid: A Study Based on an ERP Case Study

Hans-Jürgen Scheruhn[1], Daniel Ackermann[1], Roman Braun[1], and Ulrich Förster[2]

[1] Business Informatics, Media Informatics, Hochschule Harz, Friedrichstr. 57-59, 38855 Wernigerode, Germany
{dackermann,hscheruhn,u27077}@hs-harz.de
[2] Atos Werner-von-Siemens-Platz 1, 30880 Laatzen, Deutschland
ulrich.foerster@atos.net

Abstract. ERP case studies have an important impact on the transfer of knowledge between software vendor, lecturer and user. This paper describes to which extent ERP case studies can be used to demonstrate a repository based integrated approach for modeling and implementing the entire information pyramid in the context of the Business Process Management life cycle. The study is based on the Global Bike Inc. enterprise model repository created with ARIS Business Designer for SAP by Software AG. The information models have been exemplarily synchronized with SAP Solution Manager repository and implemented with SAP ERP respectively with SAP Netweaver BI.

Keywords: ERP case study, ARIS Business Designer, Business Process Management life cycle, SAP Solution Manager, SAP Netweaver BI, Sustainabilty Balanced Scorecard.

1 Problem Identification

ERP case studies are used for the transfer of knowledge in both the academic and the non-academic sector [17]. The authors state a lack of integration of strategic, tactical and operational business information models (entire Information Pyramid) as well as a lack of integration of its implementation with corresponding IT-Systems (entire business process life cycle) based on ERP case studies [21].

Scheruhn et al. [24] compared three different ERP case studies on the as-is status with each other regarding information model structures, information model types and information object types of the corresponding enterprise models [5; 30]. The comparison based on the GBI enterprise information model repository resulted in a common ERP case study framework which was worked out structurally in an evaluation matrix on the operational level.

A study about the possible enterprise information model structure of an integrated model based implementation of ERP case studies (to-be status / extended by future business objectives of tactical and strategic level) with the intention to create a common framework as repository for structured transfer of knowledge and for support of model based implementation in both the academic and the non-academic sector is yet missing.

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 446–455, 2013.

2 Objective of a Solution

ERP case studies support students learning the handling of different functional areas in a practical way [29]. Several ERP case studies exist, for example Global Bike Inc. [32], SSB Inc. [8], SAP TERP10 exercises at IDES AG [29], SAP Business ByDesign´s ALMIKA Holding [14; 15] and OPM´s IDES AG [23] which provide an introduction to different ERP systems of SAP.

This paper aims to investigate the information model based teaching of implementing ERP case studies (to-be status) in approach to further benefits. Consecutively existing ERP case studies (as-is status) are to be analyzed by students of business administration and/or business informatics as a basis for possible improvements (to-be status) and its implementation in the context of the business process life cycle [23].

The paper identifies all necessary information models based on an extended evaluation matrix created by the authors which can be used as a framework for further case study enterprise model or company enterprise model comparisons (e.g. ALMIKA Holding or VW Financial Service [31]). The integration of aspects of business and information technologies, of strategic, tactical and operational points of views as well as the integration of "actors", "business processes, business services and information flow" up to its implementation with ERP and BI systems in the context of the information pyramid in one single enterprise model is to be demonstrated [25].

3 Design and Development

3.1 Constructs

Software Reference models like the SAP Solution Manager Business Blueprint Repository [13] and/or Industry Reference models like Handels-H [2] provide support for construction processes of an individual enterprise model [5]. The enterprise models developed and evaluated in this paper are considered as integrated software-industry reference models which – in this case - depict typical company processes of the engineering industry based on software, e.g. SAP ERP. The IPR (Industry.PerformanceREADY) model is a good example [27] for such an integrated reference model approach which deals very successful in the non-academic sector. Compared to the GBI (to-be) enterprise model it covers most of the information models and objects of the evaluation matrix (see 5.) and additionally is connected with best practice KPI provided by APQC[1].

According to Scheruhn (Fig.1), the Business Process Management (BPM) life cycle encompasses five phases [23]. The main focus in this paper will be the integration of phases 2 and 3 whereas the previous paper [24] investigated phases 2 and 4. Phases 2-5 are supported (but not fully integrated) by the SAP Solution Manager and SAP ERP. SAP BI covers all phases with data seen as important constructs of processes.

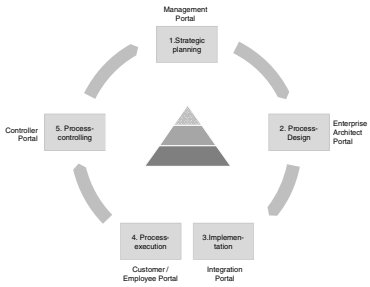


Fig. 1. Process life cycle with five phases by Scheruhn and Information Pyramid

3.2 Information Pyramid

The information pyramid is divided into three levels (Fig.1). The tactical level connects the strategic and the operational level [9]. The considered ERP case studies mainly cover the operational level in the context of the horizontal and vertical integration. In this study the tactical and strategic level on top also need to be considered. For implementing the corresponding tactical and strategic information models the authors additionally evaluated SAP Netweaver BI.

3.3 Four IT Integration Layers

In this research business processes are regarded as sub models of a model system, which includes: Subject-oriented modeling [26], Process-oriented modeling, Service-oriented modeling [3], Enterprise data modeling [6].

This leads to four horizontal IT integration layers (1.Presentation, 2.Processes, 3.Functions and 4.Data) of the enterprise model (see Fig. 2) describing an IT system like SAP ERP or SAP Netweaver BI in its interaction with a specific enterprise. A vertical decomposition of each of the four IT integration layers leads to a model hierarchy of four additional levels shown in figure 2. These hierarchical levels can be mapped with the four process levels of the SCOR reference model [28] as well as partly with the three levels of the SAP Solution Manager Business Blueprint Architecture [13].

3.4 Methods

First a hypothesis is formulated which has to be reviewed in the course of this investigation [12] based on the implementation of the entire information pyramid exemplified with SAP ERP and Netweaver BI regarding the corresponding GBI enterprise model.

Hypothesis: The available GBI ARIS to-be enterprise model can be fully integrated into SAP ERP and SAP Netweaver BI including all necessary information models to form a common information model framework repository for structured transfer of knowledge and for support of model based implementation in both the academic and the non-academic sector.

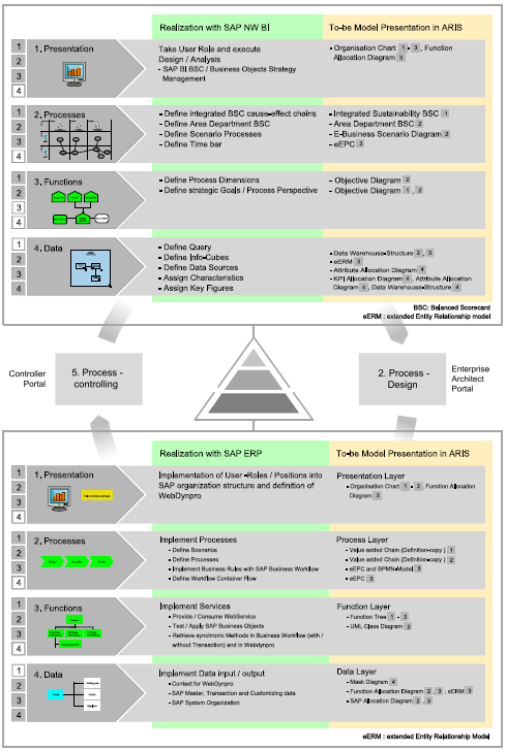


Fig. 2. GBI enterprise model for implementing SAP NW BI and for implementing SAP ERP [25]

Thus, it has to be proven to which extent SAP ERP and SAP Netweaver BI are able to interchange information model structures, information model and object types with ARIS Business Designer (BD) respectively SAP Solution Manager. These requirements will be worked out and compared in an evaluation matrix based on the following assumptions:

- SAP ERP and SAP NW BI applied together cover the entire Information pyramid and the entire Process life cycle
- SAP NW BPM is not considered because of focus on Core application without automization of Business Processes spanning different applications
- Single information models (phase two of Process life cycle) of the GBI enterprise model repository are provided completely but separately by SAP ERP and SAP NW BI
- Single information models of the GBI enterprise model repository based on ARIS Business Designer (BD) are entirely integrated in one enterprise model
- Information models of GBI based on ARIS BD and based on SAP ERP respectively SAP NW BI are equal but not directly integrated
- SAP Solution Manager serves as an interface between the information models of ARIS BD and SAP ERP respectively SAP NW BI

Area		MM	MM	MM	PP	SD	HCM
		Purchase Order	Incoming Invoice	Goods Received	Production Order	Customer Order	Business Event
Business Object		BUS2012	BUS2081	MKPF	BUS2005	BUS2032	PDOTYPE_E
Success Perspective	Objectives	Reduce Purchase Order Costs	Reduce Costs for Payment Target Overdraft	Decrease Costs for Raw Material	Decrease Production Costs	Increase Turnover	Reduce Costs fo
	KPI	Purchase Oder Costs as-is/to-be	Overdraft Costs as-is/ to-be	Costs of Raw Material as-is/to-be	Production Costs as-is / to-be	Turnover Growth as-is / to-be	Training Costs as be
	Target Ratio	5 % less costs (than Best Practice)	5 % less costs (than Best Practice)	5 % less costs (than Best Practice)	5 % less costs (than Best Practice)	10 % more turnover (than previous year)	5 % less costs (t Practice)
Customer Perspective	Objectives	Increase Internal Delivery Reliability (MM)	Increase Internal Delivery Reliability (MM)	Increase Internal Delivery Reliability (MM / Warehouse)	Increase Internal Delivery Reliability (PP)	Increase Customer Satisfaction	Increase Internal Reliability (HCM)
	KPI	Delivery Reliability as-is/to-be	Delivery Reliability as-is/to-be	Delivery Reliability (Warehouse) as-is/to-be	Delivery Reliability as-is/to-be	Customer Satisfaction as-is / to-be	Delivery Reliabili as-is/to-be
Process Perspective	Objectives	Improve Purchase Order Process	Improve Incoming Invoice Process	Improve Goods Receip t Process	Improve Production Order Process	Improve Customer Order Process	Optimize Busine Process
	KPI	Purchase Order Process cycle time	Incoming Invoice Process cycle time	Goods Received Process cycle time	Production Order Process cycle time	Customer Order Process cycle time	Event set-up cyc
	KPI	QM Code Purchase Order	QM Code Incoming Invoice	QM Code Goods Receipt	QM Code Production Order	QM Code Customer Order	QM Code Busine
Potential Perspective	Objectives	Expand MM Workflow	Expand MM Workflow	Expand MM Workflow	Expand PP Workflow	Expand SD Workflow	Expand HCM WF
	Objectives	Increase MM Skills	Increase MM Skills	Increase MM Skills	Increase PP Skills	Increase SD Skills	Increase HCM Sk
	KPI	Number of Workflows as-is/to-be	Number of Workflows as-is/to-be	Number of Workflows as-is/to-be	Number of Workflows as-is / to-be	Number of Workflows as-is / to-be	Number of Workf is / to-be
	KPI	Employee Skill Sets as-is / to-be	Employee Skill Sets as-is / to-be	Employee Skill Sets as-is / to-be	Employee Skill Sets as-is / to-be	Employee Skill Sets as-is / to-be	Employee Skill S is / to-be

Fig. 3. Area Department Balanced Scorecards and corresponding SAP Business Objects

3.5 Instantiation

The GBI (Global Bike Inc) model company is situated in the manufacturing industry. It is a multinational enterprise producing and distributing bicycles. GBI is an international group with subsidiaries in Dallas, Miami and San Diego in the United States of America and Heidelberg and Hamburg in Germany. The GBI enterprise model and its implementation have been taught at the Hochschule Harz in a degree course of Business Informatics in 2012/2013 and its results have been documented by the students in a weblog [11], which contents can be found in Chapter 4. Relevant SAP ERP models used in this enterprise model are Materials Management (MM), Production Planning (PP), Sales and Distribution (SD), Human Capital Management (HCM) and Financial Accounting (FI). GBI 2.1 case study is hosted by SAP University Alliance [18]. SAP Solution Manager 7.1 as well as SAP ERP ECC 6.0 and SAP NW BI 7.0 (7.3 in progress) are hosted by SAP UCC Magdeburg [19]. As one of the most recognized performance measurement frameworks [20], the balanced scorecard is implemented on the strategic level based on the establishment of cause and effect chains between strategic objectives respectively key performance indicators. To improve the enterprise performance in all dimensions of sustainability (economic, ecologic and social) the pillars of sustainability are integrated in the successful realization of the enterprise strategy [4]. All department Balanced Scorecards of the GBI repository are connected via internal customer relationships [11] and organized as cost centers (except sales) pursuing department IT-strategies (expand implementation of workflows). To achieve a successful realization of the department balanced scorecards (economic pillar), business rules are implemented in terms of Quality Management codes [10]. The Compliance of these business rules and a constant

comparison with best practices can be used as a key performance indicator in internal processes. This is done by a 100% integration of the strategic (top of fig. 2) with the operational models (bottom of fig. 2).

4 Demonstration

This section presents results of the first demonstration of above-mentioned ideas to implement an enterprise model repository based on the GBI model company. The corresponding blog "Online Process Management at GBI" has been posted in the SAP Community Network and currently consists of 4 (5 and 6 in progress) parts, which address following aspects of the enterprise model:

Online Process Management at GBI / Part 1 [11]	GBI Business Process Management Life Cycle, GBI Enterprise Model, GBI Sustainability Balanced Scorecard
Online Process Management at GBI / Part 2	Harz Roadmap at SAP Solution Manager, GBI Business Blueprint goes ARIS
Online Process Management at GBI / Part 3	Business Workflow and NW BPM as a major step towards process automation at GBI, How the Workflow works, Workflow Builder challenge
Online Process Management at GBI / Part 4	WebDynpro goes iPhone, Responsive Frameworks at GBI
Online Process Management at GBI / Part 5 (in progress)	Providing and consuming WebServices at GBI
Online Process Management at GBI / Part 6 (in progress)	BI at GBI

Beginning with the WebDypro Context Diagram for future business the presentation design is capable of reacting to trends like mobile and ubiquitous computing.

For WebDynpro there are possibilities to design themes with stand-alone software, for instance with an Eclipse-Plugin (ThemeEditor32). On small high resolution displays information presented with "Portal Theme Editor" themes can result in a poor readability because of pixelation and small font sizes. It should be possible to implement a basic design reacting to different browser sizes.

This could be reached by taking advantage of technologies like responsive frameworks (e.g."Twitter Bootstrap" [33]), which is based on CSS3, Html5 and Javascript. At the present those frameworks are capable of "reacting" to user devices from smartphones up to large desktop by delivering the correct font size and layout. This responsiveness is an approach of web design providing an optimal viewing experience.

Like in Wroblewski [34] and Marcotte [35] described, the responsive design method scales the html content for a wide range of displays in a very ergonomic way. The integration of basic responsive technologies in WebDynpro could lead to even more user acceptance in terms of readability and overview. As a general example Figure 4 shows the same content on 3 different devices.

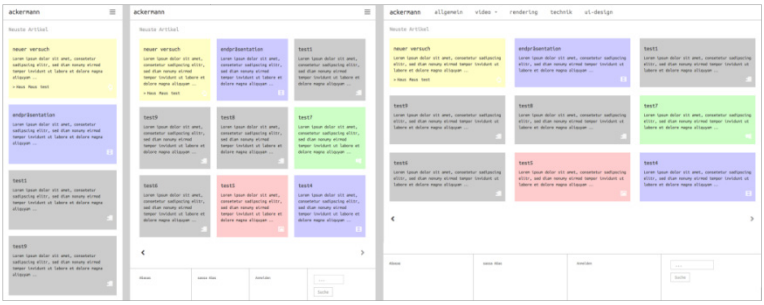


Fig. 4. Responsive layout presented on 3 different displays

5 Evaluation Matrix

An examination of the Enterprise Meta-Model (EMM) by Gudas [5] shows 8 different core objects (Actor, Process, Function, Information, Business Rule, Event, Objective and Material Flow), of which SAP Solution Manager is able to implement the first four, one per IT integration layer(fig.2). Thus, the Solution Manager offers a permeability of 50 % of object types considering the Enterprise Meta-Model by Gudas. The following Evaluation Matrix extends the EMM with reference to information model and information object types and its instances found in SAP ERP (compare bottom of Fig. 2) and SAP BI (compare top of Fig. 2) that are necessary for implementing the to-be status of GBI.

A comparison of the integrated GBI ARIS repository in its to-be status with its distributed information model repository based on SAP ERP and BI shows 80/30 % distribution (overlapping of data layer). If combined, they are able to display all aspects of the information pyramid (SAP ERP on the operational level, SAP Netweaver BI on the tactical and strategic level) and all (ERP four, BI five) of the process life cycles by Scheruhn. All the four IT-integration layers as well as the four levels of hierarchy can be integrated into one ARIS BD repository. In summary it can be said:

- 8 (EMM) + 3 (GUI, Money, Location) object types found in SAP ERP (80%) and SAP BI (30%) with 10 % overlapping of data layer = 11 object types
- All 11 object types integrated (related to each other) in 14 model types in one single GBI enterprise model in ARIS
- Only 50 % of object types (in both EMM and EMM+) and 30% of model types (EMM+) of ARIS integrated with SAP ERP / BI because of SAP Solution Manager restrictions:

- Business Blueprint not connected with SAP BI: (no Objectives, KPI)
- Business Blueprint covers only 3 of 4 hierarchy levels: (no Events, Business Rules, Money flow/material flow of ERP)
- Business Blueprint covers no object instances (e.g. no locations of enterprise)

Table 1. Evaluation Matrix GBI (to-be)

Enterprise Models		GBI to-be [ARIS BD]	GBI [SAP Solution Manager]	SAP ERP	SAP NW BI
Model structure	Level of information pyramid	1-3	3	3	1,2
	Process life cycle phases by Scheruhn	2,4	2-5	2-5	1-5
	4 IT integration layers	●	●	●	●
	Levels of hierarchy	4	3	4	4
Model types instances	Shown in Figure 2	● (1,0)	◔ (0,3)	◕ (0,8)	◔ (0,3)
Object types instances	Ref. to Process Governance Matrix by Scheruhn [22]	● (1,0)	◑ (0,5)	◕ (0,8)	◑ (0,5)
Model type instances/ 1.Presentation layer	Organization Chart	●	◑	●	○
	Function Role Allocation Diagram	●	●	●	○
2.Process layer	E-Business Scenario Diagram	●	◑	●	○
	Company and Department Balanced Scorecards	●	○	○	●
	Business Workflow of Business Rules eEPCs / BPMN Process Diagrams	●	◑	●	○
	Workflow Container Flow	●	○	●	○
3.Function layer	Business Services Tree	●	●	●	○
	Objective Diagram	●	○	●	○
	UML Class Diagram	●	○	●	○
4.Data layer	Info Cube, Dimension and Facts Data as Data Warehouse Structure Diagrams	●	○	○	●
	Data Source Model	●	○	●	●
	KPI Allocation Diagram	●	○	○	●
	System Organization Customizing Diagram	●	◑	●	○
	Mask Diagram and WebDynpro Context	●	◑	●	○
Object type instances/ 1.Presentation layer	Roles, Positions, Jobs	●	●	●	○
2.Process layer	Strategy, Perspective, Strategic Objectives , KPI	●	○	○	●
	Scenarios, Business Processes	●	●	●	○
3.Function layer	Process Steps as Methods of SAP Business Objects and External WebService	●	◔	●	○
4.Data layer	Master, Transaction and Customizing Data	●	●	●	●
	Masks (GUI)	●	●	●	○
	System Organizational Units (Locations)	●	◑	●	◑
	Keys, Foreign Keys, Describing Attributes	●	○	●	●
	Key Figures, Characteristics	●	○	○	●
	Money	●	◔	●	●
	Material	●	◕	●	●
Degree of Integration	Integration Model / Object Types seen from ARIS Business Designer	100%	30 / 50%	0 %	0%
	Integration seen from SAP Solution Manager	-	100%	100%	0%

Legend : ◔ : available only via navigation into SAP ERP
◑ : no interconnection of objects (org structure, Flow of money/material, GUI I/O)
◕ : incomplete (no persons, no instantiation)

6 Conclusion and Outlook

The evaluation matrix identifies all necessary information models on the strategic, tactical and operational level to create a common framework as a repository for structured transfer of knowledge and for support of model based implementation in both the academic and the non-academic sector. First, all of the necessary model and object type instances found in SAP ERP and SAP Netweaver BI can be (partly manually) represented and integrated in one single enterprise model repository based on ARIS

BD. Second, as they have to be imported from the ARIS BD through SAP Solution Manager, which serves as a highly integrated and powerful interface, only a fraction of the models and objects can be implemented. Seen from ARIS BD only about 30 % of the considered model type instances and only about 50 % of the object type instances can be synchronized with SAP Solution Manager, with no possibility to connect strategic or tactical model and object types. This fact is showing a need for improvement of integration of enterprise model repositories. The matrix therefore can be used for possible future improvements to show necessary steps in the creation of repository frameworks.

All information available of the GBI case study has been considered in the enterprise model. Concerning the non-academic sector it has to be noticed that the TOGAF [30] enterprise architecture reference model possibly covers more objects of a real enterprise architecture. Actually, the GBI enterprise model covers some aspects of Architecture Principles, 100% of Business Architecture, about 70 % of Information System Architecture, no Technology Architecture and about 50 % of Architecture Realization.

"Considering the implementation of SAP projects at external customers of ATOS, it can be said, that the ERP system with its processes and master data (level 2 and 4) plays the dominant role. A mapping in ARIS, besides the positive effect of visualization may also represent a chance to already show possible future extensions of the Solution Manager." [36] The authors strongly believe that the enterprise information model based on ARIS BD also can help to indicate SAP HANA database relevant areas in SAP ERP and SAP BI since most foreign keys columns are integrated.

References

1. American Productivity & Quality Center, Best Practice Knowledge Base (2013), http://www.apqc.org/APQC-knowledge-base
2. Becker, J., Schütte, R.: Handelsinformationssysteme, Redline, Frankfurt (2004)
3. Becker, A.: Nutzenpotentiale und Herausforderungen Serviceorientierter Architekturen. Gabler Verlag, Wiesbaden (2011)
4. Hahn, T., Wagner, M.: Sustainability Balanced Scorecard. von der Theorie zur Umsetzung, Universität Lüneburg (2001)
5. Gudas, S., Lopata, A.: Approach to Enterprise Modelling for Information Systems Engineering (2005)
6. Houy, C., Fettke, P., Loos, P., van der Aalst, W., Krogstie, J.: Geschäftsprozessmanagement im Großen in Wirtschaftsiformatik 06/11. Gabler Verlag (2011)
7. Hufgard, A., Krüger, S.: SAP Business by Design: Geschäftsprozesse, Technologie und Implementierung anschaulich erklärt: Galileo Press, Bonn (2012b)
8. Magal, S., Word, J.: Essentials of Business Processes and Information Systems. Wiley & Sons (2009)
9. Mertens, P., Bodendorf, F., König, W., Picot, A., Schumann, M.: Grundzüge der Wirtschaftsinformatik. Springer, Berlin (2005)
10. Nüttgens, M. et al: Governance and Sustainability in Information Systems. Managing the Transfer and Diffusion of IT. In: IFIP WG 8.6 International Working Conference, Hamburg, (2011)
11. Online Process Management at GBI / part 1 to part 4 http://scn.sap.com/community/uac/blog/2012/10/29/online-process-management-at-gbi-part-1

12. Peffers, K.: A Design Science Research Methodology for Information Systems Research. Journal of Management Information Systems 24(3), 45–77 (2007)
13. SAP Bibliothek: Business Blueprint (2012a), http://help.sap.com/saphelp_sm32/helpdata/de/2a/62c33af63ae93ae10000000a11402f/content.htm
14. SAP: Business ByDesign – Referenzsysteme - Szenario-Überblick Strategische Bezugsquellenfindung, Rel. 3.5.ss (2012a)
15. SAP: Business ByDesign: Unternehmensprofil für ALMIKA Vers.3.5. (2012b)
16. SAP: Strategische Bezugsquellenfindung - Referenzsystem für SAP Business ByDesign, Rel. 3.5. (2012c)
17. SAP:TERP10 (2012d), http://www.sap.com/germany/services/education/globaltabbedcourse.epx?context=[[|terp10|||||]]|
18. SAP UCC Magdeburg (2013), http://www.sap-ucc.com/
19. SAP University Alliance: GBI 2.1 ARIS Repository (to-be status) (2013), http://sap-schulungen.hs-harz.de/GBI_21_Sollzustand_en/index.htm
20. Schaltegger, S., Bennett, M., Burritt, R.: Sustainability Accounting and Reporting. Springer, Dordrecht (2006)
21. Scheer, A.-W., Habermann, F.: Making ERP a Success. Communications of the ACM 43(4), 57–61 (2000)
22. Scheruhn, H.-J.: Model Governance Matrix: ALMIKA (2012c), http://hscheruhn.hs-harz.de/lehre/alle/Modell-Matrix_horizontale_Navigation.pdf
23. Scheruhn, H.-J.: Online-Prozessmanagement mit SAP ERP. In: Scheruhn, H.-J., Papenfuß, D., Funk, B., Niemeyer, P. (eds.) Modellierung und Implementierung von Geschäftsprozessen in verteilten Systemen, Dr. Kovac, pp. 141–177 (2010)
24. Scheruhn, H.-J., Sicorello, S., Weidner, S.: Repository-based ERP case studies: A study about chances and benefits of agile case study development. In: Abramowicz, W., Domingue, J., Węcel, K. (eds.) BIS 2012 Workshops. LNBIP, vol. 127, pp. 186–197. Springer, Heidelberg (2012)
25. Scheruhn, H.-J., Thielert, B., Weidner, S.: GBI-Unternehmensmodell/ Business Process Analysis/Soll-Zustand (2012b), http://hscheruhn.hs-harz.de/lehre/alle/Intro_ERP_Using_GBI_Slides_BPA_de_v2.1_Sollzustand_ARIS_final.pdf
26. Schmidt, W., Fleischmann, A., Gilbert, O.: Subjektorientierter Ansatz in Prozessmanagement: Praxis der Wirtschaftsinformatik. Heft 266 4/09 (2009)
27. Software AG: Industry.PerformanceREADY (2013), http://www.softwareag.com/de/service/sap_consulting/ipr/default.asp
28. Supply Chain Council: SCOR-Referenzmodell (2012), http://supply-chain.org/scor
29. TERP10: SAP ERP - Integration of Business Processes (2012), http://www.terp10.de/pages/deutsch/kurs.php
30. TOGAF: Module 7 TOGAF Content Metamodel (2013), http://www.togaf.org/togafSlides91/TOGAF-V91-M7-Metamodel.pdf
31. Volkswagen Financial Services AG (2013), http://www.vwfsag.de/content/sites/vwcorporate/vwfsag_de/de/home.html
32. Weidner, S., Magal, S., Word, J.: Global Bike Inc. Version 2.1. (2011)
33. Twitter Bootstrap, http://twitter.github.com/bootstrap/
34. Wroblewski, L.: Mobile First. A Book Apart, New York, New York, USA (2011)
35. Marcotte, E.: Responsive Web Design. A Book Apart, New York, New York, USA (2011)
36. Förster, U: Interview at ATOS Hannover (January 28, 2013)

Agility Based on Stakeholder Interaction – Blending Organizational Learning with Interactive BPM

Christian Stary[1], Werner Schmidt[2], and Albert Fleischmann[3]

[1] University of Linz, Freistädterstraße 315, A-4040 Linz, Austria
Christian.Stary@jku.at
[2] University of Applied Sciences Ingolstadt, Esplanade 10, D-85049 Ingolstadt, Germany
Werner.Schmidt@haw-ingolstadt.de
[3] Metasonic AG, Münchner Str. 29, D-85276 Pfaffenhofen, Germany
Albert.Fleischmann@metasonic.de

Abstract. Highly dynamic adaptation of interactive work procedures not only requires structuring mechanisms, but also engaging stakeholders. As it touches business operation, not only user interface designs are challenged, but rather underlying business logic and data management issues. An inclusive perspective and interactive development support can be provided by Subject-oriented Business Process Management (S-BPM). It enables stakeholder involvement throughout an entire engineering cycle based on interaction specifications of all involved stakeholders. In this paper we propose to use S-BPM representations for both, capturing the business processes that are modified (i.e. content), and the process of modifying them along Organizational Learning (OL) cycles (i.e. adaptation). Since validated S-BPM models can be executed automatically from each stakeholder perspective, such an integrated BPM-OL approach allows structuring and implementing agility in operation under direct control of stakeholders.

Keywords: Agility, Organizational Learning Framework, Organizational Memory, Subject-oriented Business Process Management.

1 Introduction

Being forced to structural flexibility (agility) either by product or service management, customer or partner demands, organizations enter the stage of learning, both on the level of individuals and beyond [cf. 3]. Stakeholders play an important role for triggering and performing learning processes [cf. 7]. Ideally, they could drive change processes on both, the level of individual and organizational learning [cf. 2]. Subject-oriented Business Process Management (S-BPM) [cf. 4] relies on stakeholders and their capabilities constructing behavior specifications. Empirical results indicate positive effects on organizational velocity [cf. 5].

Various stakeholders can profit from a seamless support of modeling and execution [cf. 1], concerning user-interface design, business-logic, and business objects: employees assigned to functional roles in business processes, organization developers

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 456–465, 2013.

while negotiating and mediating, and IT specialists when working with stakeholders. Either starting with process analysis and proceeding with modeling, validation, and execution, S-BPM models reflecting organizational change need to be guided by a reference model on a higher level to ensure valid semantic representations und mutual understanding. However, expressing modeling-relevant information in terms of intuitive representation is crucial for effective BPM, otherwise modeling problems well known from other BPM developments, such as BPMN [cf. 6], might occur.

In section 2 a reference model for effective change management stemming from OL is reviewed. In addition, S-BPM with respect to modeling activities is recaptured, revealing basic requirements for understanding the use of the S-BPM modeling approach. Both inputs are required for informed model construction and respective learning designs on lower and higher level OL. As S-BPM allows for seamless round trip engineering due to its aligned representational and execution capabilities, the stakeholder-driven Organizational Learning models can form an effective frame of reference for agile BPM. In section 3 an S-BPM implementation of the OL approach is exemplified, providing respective S-BPM representations. Further studies to blending OL with BPM are sketched in the conclusion of the paper.

2 Foundations

In the following we describe how stakeholders can trigger learning processes to reconfigure business operation. We also review Subject-Oriented Business Process Management with respect to modeling and execution as a baseline.

2.1 Organizational Learning Framework

For intertwining individual knowledge creation and collective learning processes in [9] an experiential learning cycle has been detailed (see figure 1) where individual knowledge creation serves as input to organizational learning processes. Change on the collective layer is triggered through designing processes, and followed by access, experience, and assessment on the individual level, before being shared in work communities:

- *Design*: Stakeholders express their specific view onto information structures or business processes, according to their individual experiences.
- *Implement*: The resulting information or models can be embedded into actual task situations. Interactive artefacts could enable hands-on-experience for task accomplishment (archetyping).
- *Observe:* Stakeholders observe, in particular when interactive artefacts are used, possible effects the executed tasks have on the situation and the organization.
- *Assess*: If the results fit individual expectations or meet individual demands for change, the concerned items serve as input for the learning process on the collective level. If further process refinements or modifications are required the cycle starts again.

In order to transfer the individually elicited knowledge to the organizational level several activities facilitate interaction, collective exploration, and reflection:

- *Annotating information and process models.* Basically, all created evidence for change needs to be documented by its proposer(s). It can then become available in some kind of Organizational Memory (OM) accessible for all stakeholders. The OM is the core of a support system, as it is used to store not only originally proposed structures, but also updates and versions. Each stakeholder can then express individual concerns and formulate individual inputs in a context-sensitive way, namely through annotating information. Annotations comprise comments, associations, and supplements to the created change request.
- *Deploying and sharing representations (models).* In order to share created evidence with other stakeholders, all generated information needs to be kept. Stakeholders should be able to share their findings with others. In this way individual perspectives on a process proposal can be taken by others.

Organizational Learning can be considered as a swinging pendulum between individuals and affected stakeholders, providing inputs for change and reflecting created information (structures) collectively before putting it to operation. First, individuals take the role of content providers, before stepping into the role of receptors or respondents, reflecting content and inputs on the individual or collective layer, respectively.

A step or cycle is complete, once a modified information model, content or process descriptions have been negotiated as a commonly agreed basis for organizational development, and can become effective on the collective level. Stakeholders then take the role of receptors embodying information.

Figure 1 shows the fundamental structure, activities, knowledge management system (KMS) components, and relationships. Individual stakeholders provide inputs or start discussions, leading to information they consider being of relevance to be documented. Once created, content can be refined by other stakeholders, modified or supplemented through sharing different inputs until proceeding on the collective level.

Information and process models are initially generated on an individual level. At some point, either immediately after creating or re-modeling, or after archetyping the input provider makes his/her model public. Then, others could step in and provide annotations to reflect the model collectively. Once a version is recognized by the involved stakeholders to become operational (i.e. put to production), the build time has been completed, and a novel way of running the business at hand is ready to be deployed. Build time has come to an end, run time can start. This step is indicated by ‘transfer’ transition to the Organizational Memory (OM). At that time additional implementation steps, such as additional validation with respect to stakeholder involvement, might be performed. After completing detailed implementation, the business operation actually changes, however still in an unembodied way, as new work practices need to be experienced in actual work settings realtime to become fully embodied. Hence, the roles stakeholders take in that context are still experimenter and explorer. The roles indicate the resonance space created now for all involved stakeholders to embody novel work processes or information structures that

are stored in the Organizational Memory. Once they have embodied novel work practices (indicated in the figure by the transition 'Embodiment') they enter the individual 'Creating and Reflecting' cycle on the individual level. Even in case stakeholder experience difficulties in following novel work practices they should be able to act in a constructive way in the next learning step. In this way, the framework in figure 1 addresses explicitly the interplay between individual and collective learning. Both aspects are considered essential for interactive organizational change. The model enables considering each of them from dedicated perspectives, and their interfaces along the Organizational Learning life cycle.

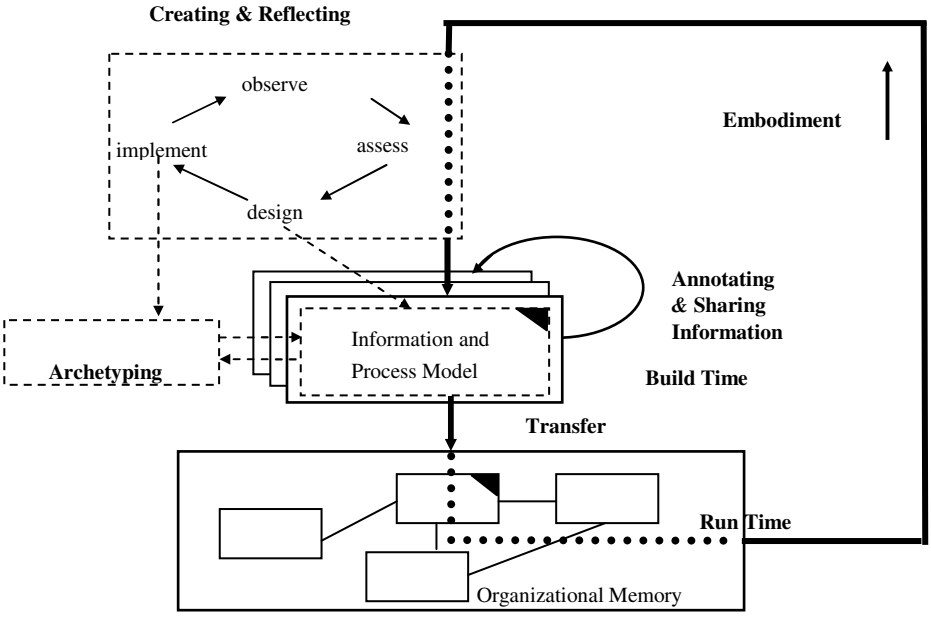


Fig. 1. The operational frame of reference and resulting KMS components for change support as given in [9]

2.2 Subject-Oriented Business Process Modeling and Execution

Subject-oriented business process models aim to bridge the gap between describing, documenting, and processing collaborative work processes. When expressing their understanding of work behavior stakeholder only have to understand interaction as the process of exchanging messages between actors and systems - which they are capable once they collaborate and use mail, in particular e-mail. Models play a crucial role in S-BPM. They are generated or processed, respectively, along a chain of various activities:

- *eliciting* and representing knowledge about work and its organization
- *analyzing* process descriptions with respect to specifity, accuracy, completeness

- *validating* models whether they could be executed in a stringent way
- *executing* models allowing hands-on experience of specifications
- *embodying* processes in organizational and technological settings of organizations
- *monitoring* with respect to expected achievements
- *simulating* in order to explore alternative or novel ways of task accomplishment

These bundles of activities represent fundamental phases of business process development. Once in each phase of development all stakeholders can participate due to their modeling capabilities, the transparency and traceability of organizational development processes can be ensured or even increased. A key enabler to this respect is the capability of support tools to execute validated process models (cf. www.metasonic.de).

We now briefly review the creation of subject-oriented representations. It starts with the identification of process-specific roles involved in the process, the subjects, and the messages exchanged between them. When sending messages, the required data is transmitted from the sender to the receiver. Thus, with a message indicating the intention to go on a business trip, e.g., 'request', sent by an employee to the supervisor, among other things the start and end date are transmitted. The behavior of communication partners, such as the supervising manager is complementary. For instance, messages sent by the employee are received by the manager, and vice versa. The manager therefore waits first in a receiving state for a business trip request from the employee. For each work procedure or business case, the

1. subjects involved in the process,
2. interactions taking place between them
3. messages they send or receive during each interaction, and
4. behavior of the individual subjects

are described as they represent the essential elements of a subject-oriented model. The description of a subject determines the order in which it sends and receives messages, and performs internal functions. Its behavior thus defines the order in which the subject processes which activities: sending or receiving, or services that are defined on the corresponding objects. Services are used to assign a specific meaning to the individual steps captured by a subject behavior model. They are triggered synchronously, i.e., a subject does not enter the corresponding next state, unless the used service has been also completely processed.

Figure 2 exemplifies a model on the left side. The screen on the right side shows a snapshot of the screen displayed when the model on the left side is executed as a workflow using the Metasonic Suite (www.metasonic.de). Once all ingredients have been added, the model can be validated and executed without further transformations. In the figure the current state is the receiving state as also highlighted in the model – an employee receives an answer to his/her request. The corresponding inbox can also be visually displayed in terms of user interface widgets as known from model-based development [cf. 8]. It requires the assignment of user-interface elements in the course of instantiating subject specifications.

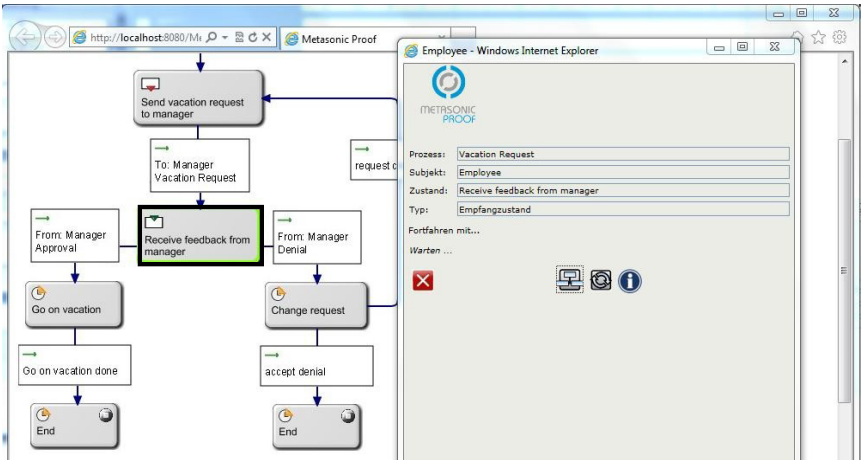


Fig. 2. Model Representation and Execution

Once those subject models have been completed that are involved in a work process, an organization-wide representation has been created. Its execution allows from each subject (i.e. stakeholder) perspective to experience the entire process interactively.

3 (Re-)Structuring Operations 'On-the-Fly'

We demonstrate how the framework given in section 2.1 can be implemented in a seamless BPM environment. Figure 3 gives the principal interaction structure of the OL process with input provider, stakeholder and organizational memory (OM) as

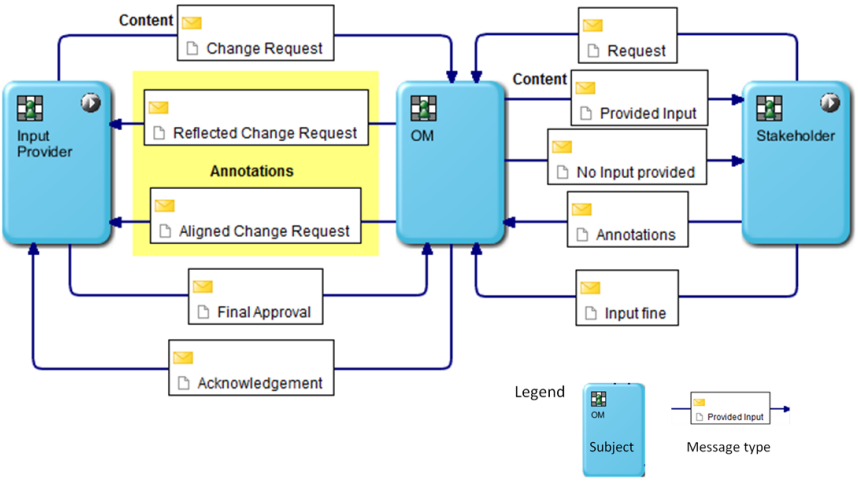


Fig. 3. Subjects involved in Organizational Learning

subjects and several types of messages they exchange. As subjects represent roles, at run time one person can act in different roles, such as an input provider and stakeholder commenting proposals for change provided by others.

For technical systems, such as the organizational memory, it is also decided at the time of implementation, which technology is going to be used at run time. In S-BPM the organizational implementation is distinguished from the technical, in order to capture both implementation aspects [4].

According to the OL life cycle (Figure 1) initially a stakeholder familiar with a work procedure (in the role of input provider) triggers change (Figure 4). He prepares some content to be considered as a new work pattern, for example a process model, such as the behavior of an employee applying for a business trip, or some other input triggering change. The content is sent to the OM as a request for change and then, the input provider waits for feedback, which may subsequently lead to changes until content has been approved by the other stakeholders. Finally the input provider embodies the new work procedure in his daily operation and acts according to it until a new learning cycle is initiated by him or other stakeholders and leads to another change of work practice.

Processing travel expense reimbursements for sales persons can serve as an example. The responsible stakeholder, e.g., an accountant, receives electronic forms containing data like name and organizational unit of the sales person, travel details, money spent etc. In order to book the reimbursement the accountant needs to retrieve the relevant cost center id of the organizational unit from a list. As this takes him 2 minutes for every single instance his idea is to let the sales people fill in the right id into the form. This would only take them some seconds because they usually know their cost center id. So the accountant provides a new version of the form including a field for the cost center id.

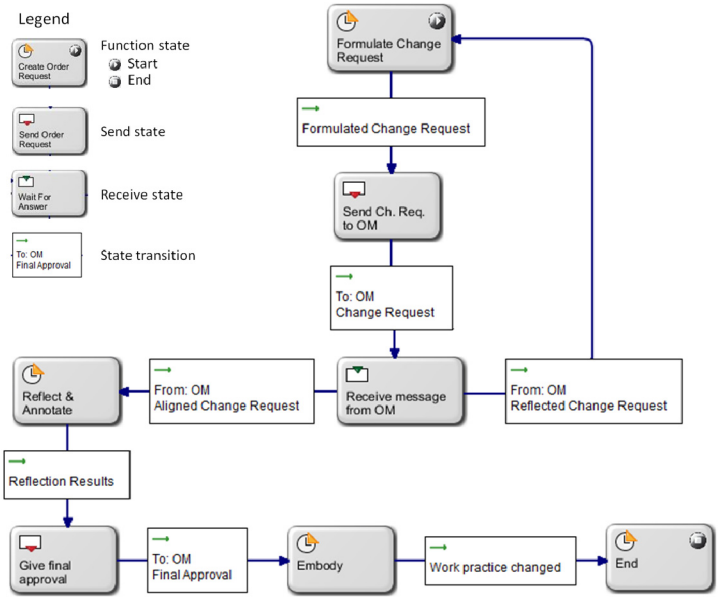


Fig. 4. Behavior of the subject 'input provider'

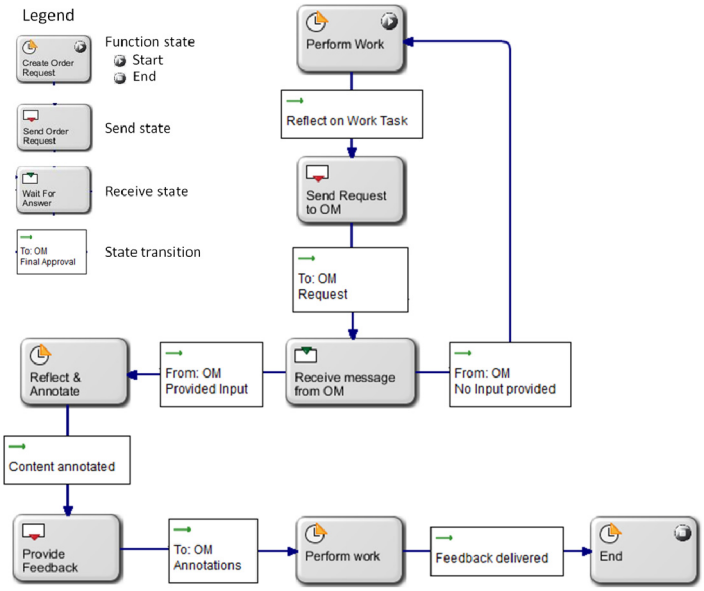


Fig. 5. Behavior of the subject 'stakeholder' responding to provided inputs

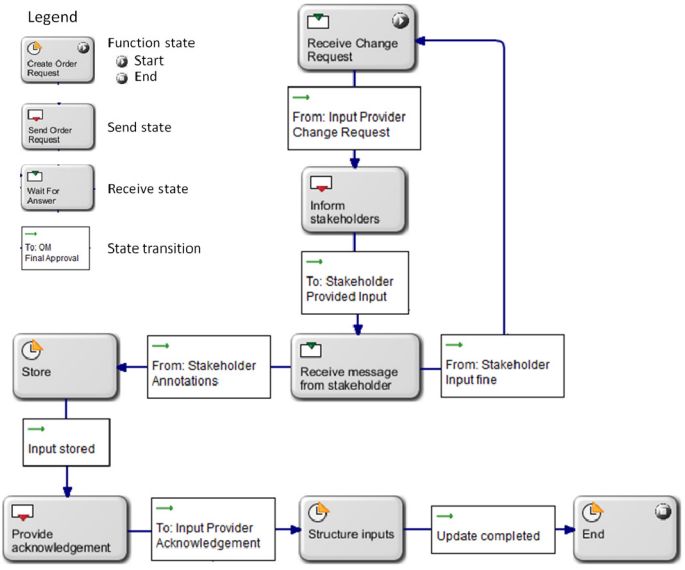


Fig. 6. Behavior of the subject Organizational Memory

Other stakeholders may, beside themselves suggesting changes to the OM, share the provided content, and make annotations (Figure 5). These annotations refer to the originally provided content, are sent to the Organizational Memory and through it become available to the input provider who may modify the change request.

The Organizational Memory receives all inputs and informs other stakeholders about the input (Figure 6). It also stores incoming annotations and provides versioning for further iterations. In our example stakeholders like sales people and the owner of the process express their acceptance of the modified form which then can become effective in operation. In another case some of them could suggest further improvement by extending the electronic form by services to automatically retrieve personal and cost center data from databases once the short employee id is entered. This could lead to an iteration, e.g., collecting the stakeholder's opinions on whether to start with the additional field or go straight for the more sophisticated change with the IT services.

The behavior diagrams have been simplified to demonstrate the interaction flow being the distinct feature of the approach. After validation it finally enables the automated execution of the behavior diagrams.

Hence, the resulting learning and business operation occurs in a choreographic way. This is of dual importance: On one hand, actors and systems may interact in parallel, however, timely synchronized. On the other hand, organizational learning is an intervention on demand, rather than prescribed, which results in non-intrusive change processes with respect to the technically informed business.

4 Conclusion

Meeting the requirement of many organizations to reconfigure their business processes dynamically, we followed a blended BPM-OL approach. A procedural framework for individual and collective changes and the operational representation for executing business processes have been intertwined, utilizing the capabilities of Subject-oriented Business Process Management (S-BPM). It provides a stakeholder perspective for modeling processes and organizational development. Re-configuring running business processes is enabled in seamless round trips.

Field studies still need to be set up to validate the results empirically in the context of BPM projects, as the current tool support is undergoing major improvements (see www.i2pm.net – Open S-BPM). Of particular importance is investigating existing role concepts, as known from BPM, e.g., process owner, and their impact on the roles identified for organizational learning, such as input provider. It is likely that mutual mappings need to be defined to perform change management on-the-fly effectively.

References

1. Avital, M., Te'eni, D.: From Generative Fit to Generative Capacity: Exploring an Emerging Dimension of Information Systems Design and Task Performance. Information Systems Journal 19(4), 345–367 (2009)
2. Böhmann, T., Burr, W., Herrmann, T., Krcmar, H. (eds.): Implementing International Services: A Tailorable Method for Market Assessment, Modularization, and Process Transfer. Gabler, Wiesbaden (2011)

3. Chen, M.T.: The Impact of Organizational Agile, Organizational Learning and Innovation on Value Creation, Ph.D., Department of Business Administration, Shih Hsin University, Taipei (2012)
4. Fleischmann, A., Schmidt, W., Stary, C., Obermeier, B.E.: Subject-oriented Business Process Management. Springer, Heidelberg (2012) http://link.springer.com/book/10.1007/978-3-642-32392-8/page/1
5. Nakamura, S., Tan, T., Hirayama, T., Kawai, H., Komiyama, S., Hosaka, S., Nakamura, M., Yuki, K.: CGAA/EES at NEC Corporation, Powered by S-BPM: The Subject-Oriented BPM Development Technique Using Top-Down Approach. In: Schmidt, W. (ed.) S-BPM ONE 2011. CCIS, vol. 213, pp. 215–231. Springer, Heidelberg (2011), doi:10.1007/978-3-642-23471-2_15
6. Recker, J.: Opportunities and Constraints: The Current Struggle with BPMN. Business Process Management Journal 16(1), 181–201 (2010)
7. Rouse, W.B. (ed.): Enterprise Transformation: Understanding and Enabling Fundamental Change. Wiley, Hoboken (2006)
8. Stary, C.: TADEUS: Seamless Development of Task-based and User-oriented Interfaces. IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans 30(5), 509–525 (2000)
9. Stary, C., Fleischmann, A.: Evidence-based Interactive Management of Change. Knowledge Management & E-Learning: An International Journal (KM&EL) 3(2), 170–200 (2011) http://www.kmel-journal.org/ojs/index.php/online-publication/article/view/31/83

Analysis of Customer Preference through Unforced Natural Passive Observation

Terumasa Tajima[1], Yusuke Iida[1], and Toshikazu Kato[2]

[1] Graduate School of Science and Engineering, Chuo University, 1-13-27 Kasuga, Bunkyo-ku, Tokyo, 112-8551 Japan
[2] Faculty of Science and Engineering, Chuo University, 1-13-27 Kasuga, Bunkyo-ku, Tokyo, 112-8551 Japan
terumasatajima0126@gmail.com

Abstract. In our former research, customer's preference has been estimated by passive observation of shopping behavior, e.g. customer's "look" and "touch". It takes much time to understand their preferences form the log. We need quickly to build up the preference model to perform suitable recommendation for a new customer. For this reason, we will propose an active observation mechanism that detects customer's unforced natural behavior to information through ambient devices such as speakers and electric displays. This mechanism also analyzes customer's preference on features and their values of commodities, which enables the system to estimate the rate of preference to an unknown product. We have experimented on ten university students. We had them evaluate the thirty-six Shirts. We used these evaluations for precision evaluations in naive Bayes classifier. We used the leave-one-out cross-validation. As the result, we have achieved the average precision in the estimating preferences by naive Bayes classifier is 71%.

Keywords: Passive and active observation, Recommendation system, Machine classifier, Digital signage, Decision making support

1 Introduction

Customer's preference has been analyzed on shopping logs in a real shop as well as page view logs in web shopping sites [1][2][3][4][5]. These approaches require mass data for each individual to build personal preference model, otherwise the analysis is just on the mass market as an average preference of customers. We have developing a smart shop system that analyzes each customer's preference to commodities observing his unforced natural behavior in the shop [6]. We have adopted passive observation of the customer's behavior, such as "look", "touch", and "take" a commodity in a shelf, which requires some period to save log. We will also adopt active observation of the customer's response to the digital signage, such as gaze and ignore, to capture log more quickly without forcing him to answer the messages.

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 466–474, 2013.

2 Discussion on Passive Observation

In our former research, customer's preference has been estimated by passive observation of shopping behavior, e.g. customers "look" or "touch". As this observation can only be captured their behavior without forcing replies, it takes much time to understand their preferences form the log. Quickly we have to build up the mechanism of preference model to perform preferable recommendation for a new customer. For this reason, we will propose on active observation mechanism that detects customer's unforced natural behavior to information through ambient devices such as speakers and electric displays. This mechanism also analyzes customer's preference for features and their values of commodities, which enables the system to estimate the rate of preference to an unknown product.

3 Active Observation

The aim of our research is to propose active observation system that can estimate customers' preferences from small load. This section explains three things. First is a way of observing behaviors that enables us to estimate customer's preferences. Second is a method to estimate customer's preferences from behaviors. Third is a method to understand customers' preference from small load. We described the overview process of active observation in Fig. 1.

3.1 A Way of Observing Behaviors

As devices of the active observation system, we adopted ambient display as digital signage devices with a function to detect a human face and its direction.

3.2 Estimating a Preference from Behaviors

Our former research found correlation between the gaze time and the rate of preference. It considers the subjectivity of gaze behavior by customer analysis [7]. The gaze time is the total duration of watching information on the ambient device. This observation system captures gaze time of each customer. Thereby, we estimate the customer's preference, "like" or "dislike", to be compare the mean and the variance of gaze time. For example, Fig 1 (a) (a') shows that the system estimates red polo shirt "like" and blue turtleneck shirt "dislike".

3.3 Analyzing Customer's Preference

We enable the system to estimate the rate of preference to unknown products. For that analyzing and learning algorithm, we have embedded content based filtering to

system. We have used the naive Bayes classifier. Thereby, we can get an evaluation for the product which a customer hasn't evaluated. For example, Fig 1 (b) (b') shows that the system computed a posterior probability with product property of polo shirt and turtleneck shirt. Thereby, it estimated rates of preference the similar products in the database. On the other hands, content based filtering is recommending only similar products which customer evaluates as high. However, the customer's over all preference cannot have been analyzing.

We explain model making and a classification by naive Bayes. The way of calculating naive Bayes probabilistic model wants you to refer to documents [8]. The probability model for a classifier is a conditional model. Conditional model is dependent class variable, Class, and feature variables, sample, by eq. (1).

$$P(Class \mid sample) \tag{1}$$

Furthermore, using Bayes' theorem, we obtain the following expression (2).

$$P(Class \mid sample) = \frac{P(Class)P(sample \mid Class)}{P(sample)} \propto P(Class)P(sample \mid Class) \tag{2}$$

The feature variable, sample, is a set of the several feature variables $word_1$ through $word_n$. Each feature $word_i$ is conditionally independent of every other feature $word_j$ for j ≠ i given the class variable. We obtain the following expression (3).

$$P(sample \mid Class) = P(word_1 \wedge ....word_i \mid Class) = \prod_i P(word_i \mid Class) \tag{3}$$

Thus, we had classified unknown samples by eq. (4).

$$Class_{map} = \arg\max_{Class} P(Class) \prod_i P(word_i \mid Class) \tag{4}$$

We have embedded a metric based on entropy to estimate rates of recommendations. For example, Fig 1 (c) (c') shows that the system computes posterior probability of property on product (P (redllike), P (redldislike)). Entropy, H, are calculated from posterior probability, P, by eq. (5).

$$H = -P\ log_2 P - (1-P)\ log_2(1-P) \tag{5}$$

Therefore, we calculated the entropy of each product and decided a next candidate for recommendation from entropy. We combine entropy with naive Bayes classifier. Thereby, we can analyze the customer's preference with short duration and a small load.

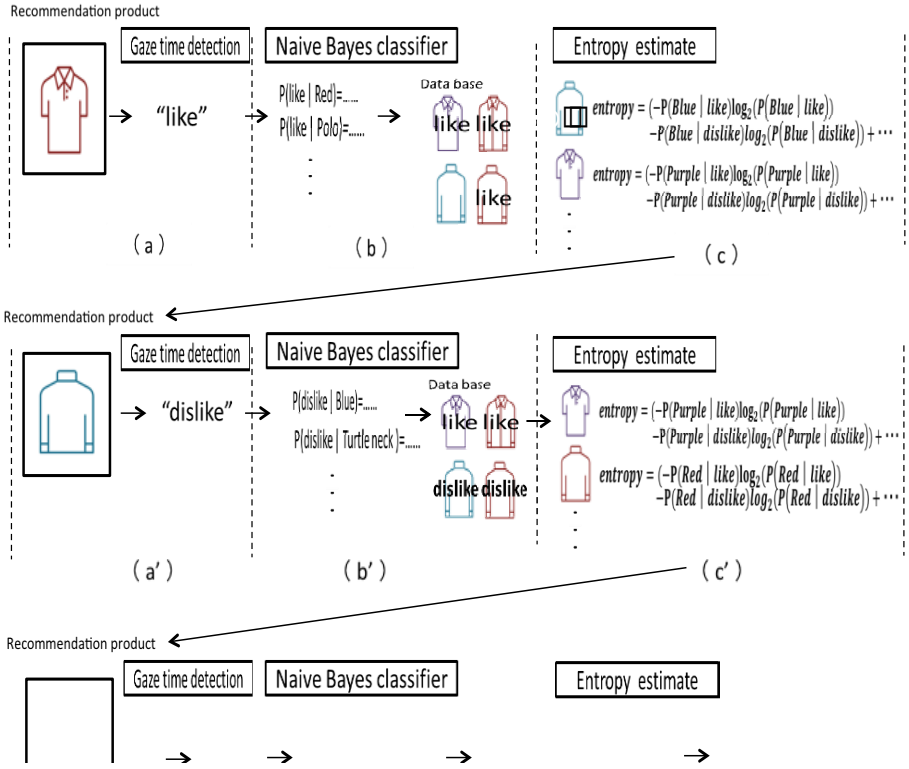


Fig. 1. Image of the flow to estimated customer's preference

(a)(a') … We estimated the customer's preference, "like" or "dislike", to compare the mean and the variance of gaze time.

(b)(b') … We computed a posterior probability with product property and estimated similar product in the database.

(c)(c') … We calculated the entropy of each product and decided a next candidate for recommend from entropy

4 Experiment

We had been experimenting on the effectiveness of the two processes. First is effectiveness of the estimated precision of the customer's preference by naive Bayes classifier. Second is effectiveness of the customer's preference estimate number of times by the entropy.

4.1 Estimating a Preference by Naive Bayes Classifier

We have experimented on ten university students (21- 24 years-old, in Tokyo). We had them look and touch thirty-six Shirts. Also, we had them evaluate the Shirts. We

used these evaluations for precision evaluations in naive Bayes classifier. We used the leave-one-out cross-validation.

4.2 Reduction of the Estimate Duration by Entropy

We have experimented on three university students (21- 24 years-old, in Tokyo).

We had them look one hundred and four Shirts. Furthermore, we had subjects evaluate one hundred and four Shirts. We used these evaluations in analysis and precision evaluations. We compared the number of times to preference estimate when using the entropy with does not use it.

5 Result

5.1 Estimating a Preference by Naive Bayes Classifier

The result is as follows (Table1).

Table 1. Precision by naive Bayes classifier

Subject No	Precision(%)
1	61.1
2	75.0
3	69.4
4	72.2
5	61.1
6	66.7
7	80.7
8	80.6
9	80.6
10	72.2
average	71.1

As the result, the average precision in the estimating preferences by naive Bayes classifier is 71%. Therefore, we think that the estimating preferences by naive Bayes classifier are effective. Hence, we think that we enable the system to estimate the rate of preference to the unknown products.

5.2 Reduction of the Estimate Duration by Entropy

First of all, we explain the figure of this chapter in Fig. 2.

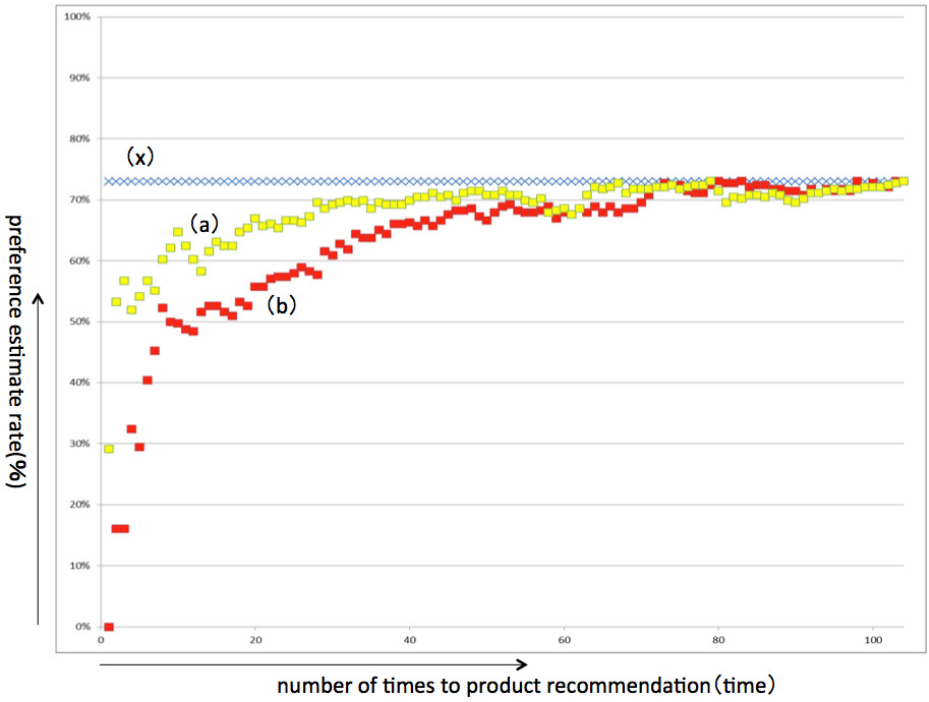


Fig. 2. Example of the figure

(x) … The upper bound that we can preference estimate in naive Bayes
(a) … Relationship between the number of times to product recommendation and preference estimate rate when using the entropy.
(b) … Relationship between the number of times to product recommendation and preference estimate rate when to not use the entropy.

The result is as follows (Fig. 3, Fig.4 and Fig. 5).

As analysis, we analyzed the change of the preference estimate rate when a product recommended by evaluation of the subject. We applied the analysis three times and averaged the results every subject.

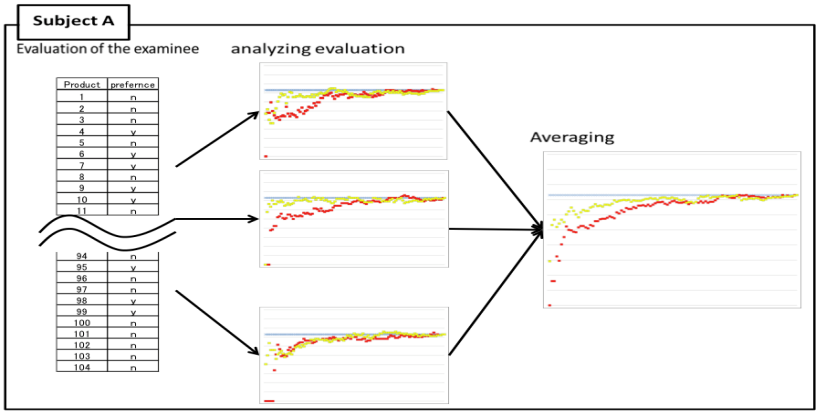


Fig. 3. Analysis result of subject A

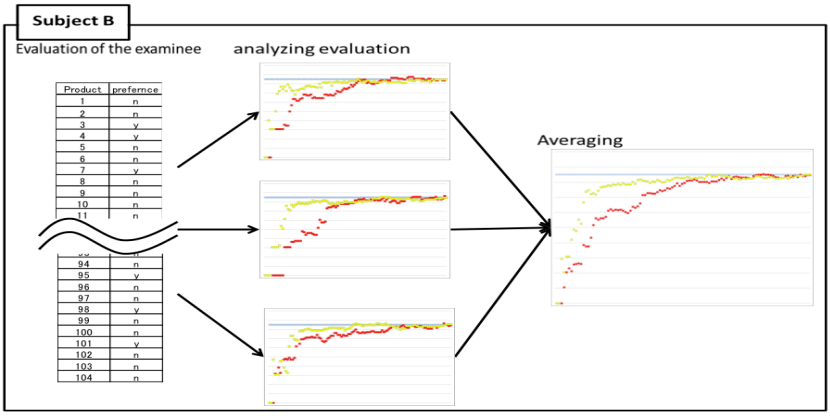


Fig. 4. Analysis result of subject B

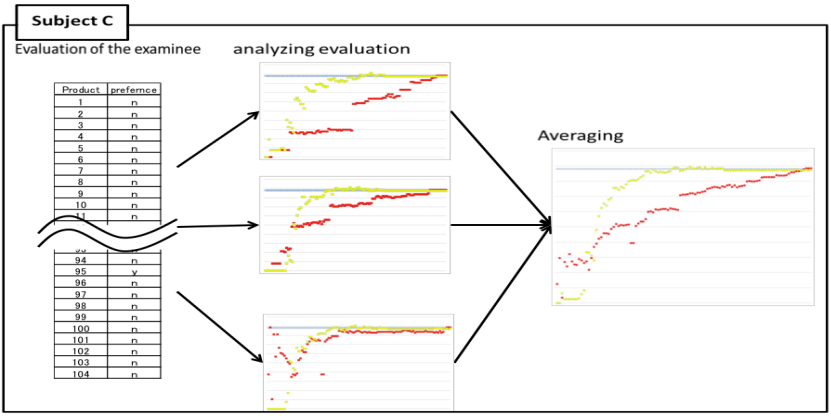


Fig. 5. Analysis result of subject C

Furthermore, we gathered up all the results of the subject in Fig. 6.

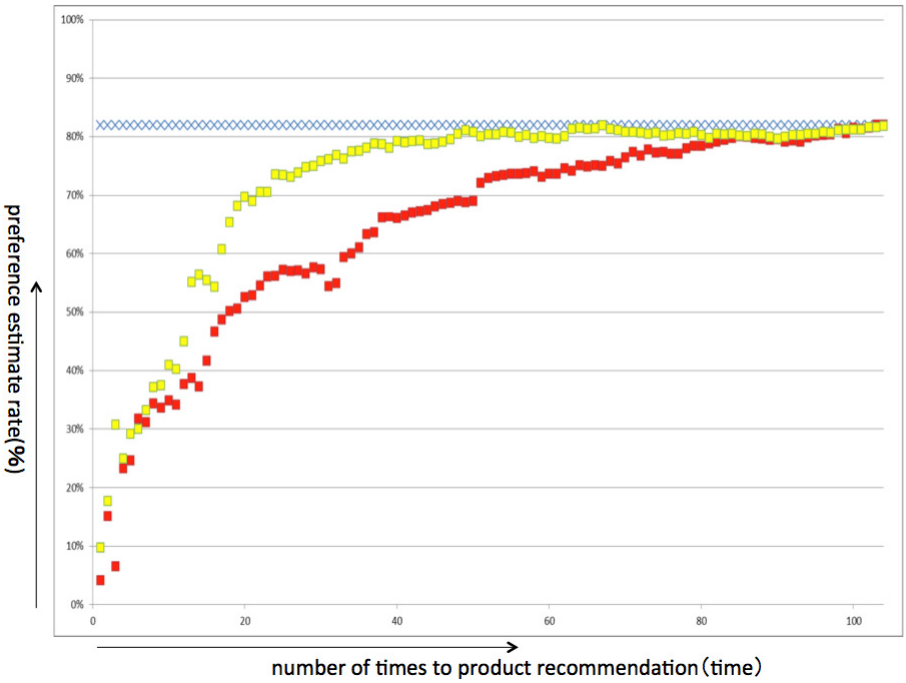


Fig. 6. Preference estimate when using the entropy with does not use it (united in subjects)

As the result, the learning using the entropy may achieve a high preference estimate rate in a short duration. As much analysis results, the learning to use the entropy is shorter than learning not to use. Hence, we think that we can analyze the customer's preference with short duration and a small load.

6 Conclusion

Our purpose has developing a smart shop system that analyzes each customer's preference to commodities observing his unforced natural behavior in the shop. In our former research, customer's preference has been estimated by passive observation of shopping behavior, e.g. customer's "look" and "touch". It takes much time to understand their preferences form the log. We need quickly to build up the preference model to perform suitable recommendation for a new customer. For this reason, we proposed an active observation mechanism that detects customer's unforced natural behavior to information through ambient devices such as speakers and electric displays. This mechanism also analyzes customer's preference on features and their values of commodities, which enables the system to estimate the rate of preference to an unknown product. We had been experimenting on the effectiveness of the two

processes. First is effectiveness of the estimated precision of the customer's preference by naive Bayes classifier. Second is effectiveness of the customer's preference estimate number of times by the entropy. As the result, we have achieved the average precision in the estimating preferences by naive Bayes classifier is 71%. We can expect to analyze that we can analyze the customer's preference with short duration and a small load by using the entropy.

Acknowledgement. This work was partially supported by JSPS KAKENHI grants, "Effective Modeling of Multimodal KANSEI Perception Processes and its Application to Environment Management" (No. 24650110), "Robotics modeling of diversity of multiple KANSEI and situation understanding in real space" (No. 19100004) and TISE Research Grant of Chuo University, "KANSEI Robotics Environment".

References

1. Fayyad, U.M., Wierse, A., Grinstein, G.G.: Information Visualization in Data Mining and Knowledge Discovery. Morgan Kaufmann (2001)
2. Eirinaki, M., Vazirgiannis, M.: Web mining for web personalization. ACM Transactions on Internet Technology (TOIT) 3(1), 1–27 (2003)
3. Yukihiro, A.: In-Store Buying Behavior and Retail Management(1). Journal of Business Administration of Kwansei Gakuin University 32(4), 117–146 (1985)
4. Yukihiro, A.: In-Store Buying Behavior and Retail Management(2). Journal of Business Administration of Kwansei Gakuin University 33(1), 163–179 (1985)
5. Nakayama, A., Tsurumi, H.: Analysis of Customer's Multiple Purchase Behaviors and Influences of Store Arrangement in Department Store by INDSCAL. Jpn.J. of Behaviormetrics 32(2), 201–208 (2005)
6. Imamura, N., Ogino, A., Kato, T.: PatternAnalysis of Customer Behavior Using Ubiquitous Interface. In: Proceedings of the Annual Conference of JSKE, vol. 8(3), pp. 741–747 (2009)
7. Iida, Y., Kato, T.: Presumption of interest by active sensing using measurement of natural reaction to recommendation. In: Proceedings of the Annual Conference of JSKE, vol. 13, p. 26 (2011)
8. Manning, C.D., Raghavan, P., Schütze, H.: Introduction to Information Retrieval. Cambridge University Press (2004)

The Development of an Innovative Design Process for Eco-efficient Green Products

Shiaw-Tsyr Uang[1] and Cheng-Li Liu[2]

[1] Department of Industrial Engineering and Management,
Minghsin University of Science & Technology, Hsinchu, Taiwan
[2] Department of Industrial Management, Vanung University, Taoyuan, Taiwan
uang@must.edu.tw

Abstract. Along with the technological advances and the rapid flow of information, people continue to create more innovative technology products in order to make life more convenience and comfortable than the past. However, huge industrial processes bring significant changes in the environment, the earth's resources are getting to dry up, the natural environment and ozonosphere have been destroyed, and leading serious threats to the survival of the Earth's species. Therefore, in recent years, environmental awareness began to be respected, and people explore green issues, hoping to make the earth resources and living things can be survival sustainable. This research attempts to integrate 39 engineering parameters of TRIZ with 7 eco-efficient elements to construct a green product design matrix, and use the patents of LED lights to verify the validity of this innovative green product design process. The green product design process of this research may allow designers to generate green products by improving eco-efficient elements as well as resolving conflict design issues, and assist organizations to increase competitiveness and to achieve sustainable development goals.

Keywords: Theory of Inventive Problem Solving (TRIZ), Eco-efficient, Green Design, Product Innovation, Light Emitting Diode (LED).

1 Introduction

Along with the technological advances and the rapid flow of information, people continue to create more innovative technology products in order to make life more convenience and comfortable than the past. Product designers usually concentrate on design issues such as product efficacy, cost, appearance, usability, and so on. However, huge industrial processes bring significant changes in the environment, the earth's resources are getting to dry up, the natural environment and ozonosphere have been destroyed, and leading serious threats to the survival of the Earth's species [7][20]. Product designers are now facing brand-new challenges than before. In recent years, environmental awareness began to be respected, and people explore green issues, hoping to make the earth resources and living things can be survival

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 475–483, 2013.

sustainable. Under these growing environmental pollutions, many countries have implemented environmental protection laws to reduce the environment impact of industry. For instance, WEEE (Waste Electrical and Electronic Equipment Directive), RoHS (Restriction of Hazardous Substances Directive) and EuP (Energy Using Product) organizations under the auspices of the European Union limits the environmental pollution caused by products [2][19]. Product designers should thus provide eco-friendly solutions to product design problem through innovative methods and tools to help implement eco-friendly designs [1][14].

As defined by the World Business Council for Sustainable Development (WBCSD) [20], "eco-efficiency is achieved by the delivery of competitively priced goods and services that satisfy human needs and bring quality of life, while progressively reducing ecological impacts and resource intensity throughout the life-cycle to a level at least in line with the Earth's estimated carrying capacity." In short, eco-efficiency is concerned with creating more product values with less negative environmental impacts.

In addition, eco-efficiency can be implemented along the entire value chain of a product or service [4][9][17]. For some companies, the main harmful effects on the environment with which they are associated actually occur outside their own organizations – either upstream in the raw material generation and supplier processing phases, or downstream in the product use or disposal phases. In light of this, WBCSD pointed out eco-efficiency can be achieved through seven key eco-efficient elements as below (REDUCES, in short) [20]. It will produce high eco-efficient products or services while improving these seven eco-efficient elements.

1. **R**educe material intensity
2. **E**nergy intensity minimized
3. **D**ispersion of toxic substances is reduced
4. **U**ndertake recycling
5. **C**apitalize on use of renewable resources
6. **E**xtend product durability
7. **S**ervice intensity is increased.

Theory of Inventive Problem Solving was developed by Genrich Altshuller and his colleagues, and is now being developed and practiced as an innovative tool throughout the world. "TRIZ" is the acronym for this theory in Russian [3][5][6]. TRIZ is a dialectic way of thinking in finding a suitable solution while facing a design predicament [10] [18].

The contradiction matrix of TRIZ provides designers which of the 40 inventive principles have been used most frequently to solve a problem that involves a particular contradiction [11]. For example, if you need a static object to be longer without becoming heavier. The improving feature of this contradiction is "04.Length of non-moving object" and the worsening feature is "02.Weight of non-moving object". As indicated in Table 1, we may use inventive principles "35.parameter changes", "28.mechanics substitution", "40.composite materials", and "29.pneumatics and hydraulics" to solve this design problem.

Table 1. Part of the TRIZ contradiction matrix

Worsening feature / Improving feature	01.Weight of moving object	02.Weight of non-moving object	…	38. Level of automation	39. Productivity
01.Weight of moving object	+	-	…	26.35.18.19.	35.03.24.37.
…	…	…	…	…	…
04.Length of non-moving object	-	35.28.40.29.	…	-	30.14.07.26.
…	…	…	…	…	…
…	…	…	…	…	…
39.Productivity	35.26.24.37.	28.27.15.3.	…	5.12.35.26.	+

The purpose of this research is to integrate engineering parameters of TRIZ with eco-efficient elements to construct a green product design process, and use the patents of LED lights to verify the validity of this green product design process.

2 Method

Research procedures are introduced in the following sections.

First, in order to transfer 39 engineering parameters of TRIZ into 7 eco-efficient elements, literature reviews [1][8][12][13][18][21][22][23] are conducted to find out the relationship between them. For example, reducing material intensity of a product may be achieved by changing product's features, such as weight, size, shape, and so on. Table 2 demonstrates the relationship between 39 engineering parameters and 7 eco-efficient elements based on literature reviews.

Next, Table 2 is used to transform/reduce the original TRIZ 39×39 contradiction matrix into 7×39 green design matrix. Since the matrix is downsized in this study, the number of suggested inventive principles increases within each cell. Hence, inventive principles are sorted descending by their frequencies within each cell of the green design matrix. And, inventive principles with frequencies less than two are eliminated from the table. Table 3 is part of the green design matrix constructed by this study. Most frequently used inventive principles are shown in the cells under various combinations of design contractions.

Table 2. Part of the relation table between engineering parameters and eco-efficient elements

Engineering parameter \ Eco-efficient element		R	E	D	U	C	E	S
1	Weight of moving object	*	*					
2	Weight of non-moving object	*						
3	Length of moving object	*	*					
4	Length of non-moving object	*						
5	Area of moving object	*	*					
6	Area of non-moving object	*						
7	Volume of moving object	*	*					
8	Volume of non-moving object	*						
9	Speed				*			*
10	Force				*			
11	Tension/ pressure				*			
12	Shape	*						
13	Stability of object			*			*	
14	Strength	*				*	*	
15	Durability of moving object						*	
							*	

Table 3. Part of the proposed green design matrix

Worsening feature / Improving eco-efficient element	01. Weight of moving object	02.Weight of non-moving object	…	38. Level of automation	39. Productivity
Reduce material intensity	29.40.08.15.35.	27.35.26.	…	35.26.	10.35.28.02.14.26. 29.34.
Energy intensity minimized	29.15.28.	27.19.	…	02.26.16.28.35.	28.35.10.02.34.
Dispersion of toxic substances is reduced	35.	35.26.	…	35.	35.
Undertake recycling	28.18.26.	28.35.13.26.	…	28.	10.28.35.
Capitalize on use of renewable	27.	02.27.	…	34.	10.35.
Extend product durability	02.08.21.27.35.39. 40.	27.	…	01.13.34.35.	35.10.
Service intensity is increased	35.10.13.26.28.	10.28.	…	35.34.	35.

The innovative design process of this research is depicted in Figure 1. Product designers may consider and select eco-efficient elements according to their priorities, and then analyze the worsening feature under this condition. After that, the improving eco-efficient element and worsening feature can be used to locate inventive principles in green design matrix. Finally, designers may adopt suggested inventive principles to design/improve products.

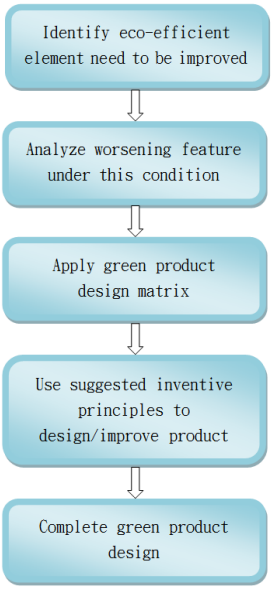


Fig. 1. Innovative design process for eco-efficient product

The validity of the above innovative green product design process is verified and case-studied by existent patents of LED lights. Detailed description of two patents is illustrated in the Results section.

3 Results and Discussion

The proposed green design matrix is examined by existent patents. Our cases used to evaluate the validity of the proposed green design model are selected from the Taiwan Patent Database.

First one (Patent No M348049) is an indicator light on the bike (Figure 2) [15]. One piezoelectric module is installed under the bike saddle. While riding a bike, the induced vibration drives the piezoelectric module to generate lighting which can be used to alert other road users.

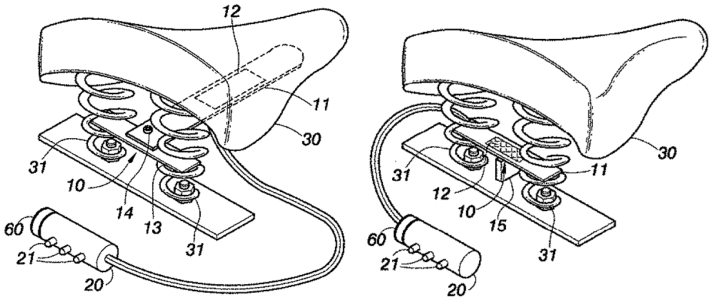


Fig. 2. Diagram of Taiwan Patent No M348049 [15]

This patent wants to minimize energy intensity; however, any imbedded devices may cause waste of substance. Hence, as shown in Table 3, green design matrix suggests inventive principles 35, 10, 18, 27, 28, 31 can be applied to solve this design problem. This patent obviously adopts inventive principle "18.mechanical vibration" as suggested.

Table 4. Part of the green design matrix used to illustrate Patent No M348049

Worsening feature / Improving eco-efficient element	01. Weight of moving object	…	23. Waste of substance	…	39. Productivity
Reduce material intensity	29.40.08.15.35.	…	10.35.39.03.05.28.	…	10.35.28.02.14.26. 29.34.
Energy intensity minimized	29.15.28.	…	35.10.18.27.28.31.	…	28.35.10.02.34.
…	…	…	…	…	…
…	…	…	…	…	…
…	…	…	…	…	…

Second one (Patent No I327632) is a multi-functional LED camping light (Figure 3) [16]. Fixtures can be easily setting to change light direction according to users' preferences.

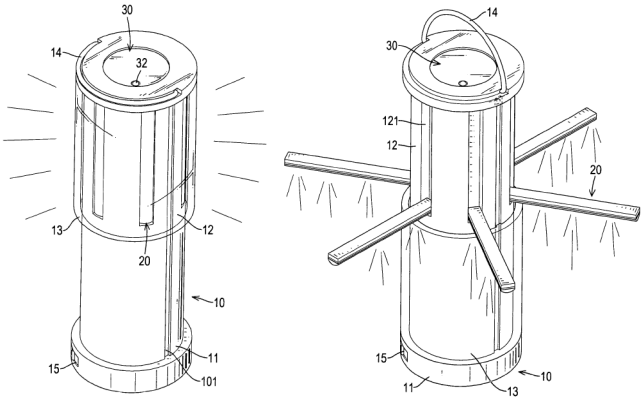


Fig. 3. Diagram of Taiwan Patent No I327632 [16]

This patent wants to reduce material intensity while camping; however, it may cause inconvenience of use. Hence, as shown in Table 4, green design matrix suggests inventive principles 02, 13, 15, 32, 16, 25, 35 can be applied to solve this design problem. This patent obviously adopts inventive principle "15.dynamicity" by adjusting setting as suggested.

Table 5. Part of the green design matrix used to illustrate Patent No I327632

Worsening feature / Improving eco-efficient element	01. Weight of moving object	…	33. Convenience of use	…	39. Productivity
Reduce material intensity	29.40.08.15.35.	…	02.13.15.32. 16.25.35.	…	10.35.28.02.14.26. 29.34.
…	…	…	…	…	…
…	…	…	…	…	…
…	…	…	…	…	…
Service intensity is increased	35.10.13.26.28.	…	01.28.34.	…	35.

Case studies of the existence patents demonstrate the validity of this green product design matrix. The proposed green design matrix may simplify the process of selecting the most appropriate inventive principle to solve a specific contradiction between eco-efficient element and engineering feature.

4 Conclusion

With raising environmental protection awareness, product designers are now facing brand-new challenges than before. Product designers should thus provide eco-friendly solutions to product design problem through innovative methods and tools to help implement eco-friendly designs. Present research attempts to integrate 39 engineering parameters of TRIZ with 7 eco-efficient elements to construct a green product design matrix, serving as an innovative green design tool. Case studies of two LED patents verify the validity of the proposed innovative green product design process. The proposed green design matrix may simplify the process of selecting the most appropriate inventive principle, and then to accelerate design problem solving. To sum up, the green product design process of this research may allow designers to generate green products by improving eco-efficient elements as well as resolving conflict design issues, and assist organizations to increase competitiveness and to achieve sustainable development goals.

Acknowledgement. The authors would like to thank the National Science Council of the Republic of China for financially supporting this work under Contract No. NSC101-2221-E-159-002.

References

1. Chang, H.-T., Chen, J.L.: The conflict-problem-solving CAD software integrating TRIZ into eco-innovation. Advances in Engineering Software 35, 553–566 (2004)
2. Chu, C.-H., Luh, Y.-P., Li, T.-C., Chen, H.: Economical green product design based on simplified computer-aided product structure variation. Computers in Industry 60, 485–500 (2009)
3. D'Anna, W., Cascini, G.: Supporting sustainable innovation through TRIZ system thinking. Procedia Engineering 9, 145–156 (2011)
4. Dangelico, R.M., Pontrandolfo, P.: From green product definitions and classifications to the Green Option Matrix. Journal of Cleaner Production 18, 1608–1628 (2010)
5. Fresner, J., Jantschgi, J., Birkel, S., Barnthaler, J., Krenn, C.: The theory of inventive problem solving (TRIZ) as option generation tool within cleaner production projects. Journal of Cleaner Production 18, 128–136 (2010)
6. Hou, J., Su, D.: A customer-manufacturer-competitor orientation model for product life cycle analysis based on QFD, AHP/ANP and TRIZ. International Journal of Design Engineering 1(1), 104–124 (2007)
7. Intergovernmental Panel on Climate Change (IPCC), http://www.ipcc.ch/
8. Jones, E., Harrison, D.: Investigating the use of TRIZ in eco-innovation. The TRIZ Journal (2000), http://www.triz-journal.com
9. Kobayashi, H.: A systematic approach to eco-innovative product design based on life cycle planning. Advanced Engineering Informatics 20, 113–125 (2006)
10. Li, S.-L., Huang, H.-H.: Applying TRIZ and fuzzy AHP to develop innovative design for automated manufacturing systems. Expert Systems with Applications 36, 8302–8312 (2009)

11. Lim, I., Sheng, S., Kok-Soo, T.: Eco-efficient product design using theory of inventive problem solving (TRIZ) principles. American Journal of Applied Sciences 7(6), 852–858 (2010)
12. Liu, C.-C., Chen, J.-L.: Development of product green innovation design method. In: Proceedings of Eco Design, pp. 168–173 (2001)
13. Liu, C.-C., Chen, J.-L.: Inventive principles and evaluation method for eco-innovative design of products. Environmentally Conscious Design and Inverse Manufacturing, 369–376 (2003)
14. Pujari, D., Wright, G., Peattie, K.: Green and competitive influences on environmental new product development performance. Journal of Business Research 56, 657–671 (2003)
15. Taiwan Patent No M348049, http://twpat2.tipo.gov.tw/tipotwoc/tipotwkm?@@1083144416
16. Taiwan Patent No I327632, http://twpat2.tipo.gov.tw/tipotwoc/tipotwkm?@@1083144416
17. Tseng, H.-E., Chang, C.-C., Li, J.-D.: Modular design to support green life-cycle engineering. Expert Systems with Applications 34, 2524–2537 (2008)
18. Uang, S.-T., Liu, C.-L., Chang, M.: A Product Design Approach by Integrating Axiomatic Design and TRIZ. In: Marcus, A. (ed.) Design, User Experience, and Usability, HCII 2011, Part II. LNCS, vol. 6770, pp. 225–233. Springer, Heidelberg (2011)
19. Vezzoli, C., Sciama, D.: Life cycle design: from general methods to product type specific guidelines and checklists: a method adopted to develop a set of guidelines/checklist handbook for the eco-efficient design of NECTA vending machines. Journal of Cleaner Production 14, 1319–1325 (2006)
20. World Business Council for Sustainable Development (WBCSD), http://www.wbcsd.org/home.aspx
21. Yang, C.J., Chen, J.L.: Accelerating preliminary eco-innovation design for products that integrates case-based reasoning and TRIZ method. Journal of Cleaner Production 19, 998–1006 (2011)
22. Yang, C.J., Chen, J.L.: Forecasting the design of eco-products by integrating TRIZ evolution patterns with CBR and simple LCA methods. Expert Systems with Applications 39, 2884–2892 (2012)
23. Zhang, J., Shang, J.: Research on developing environmental protection industry based on TRIZ theory. Procedia Environmental Sciences 2, 1326–1334 (2010)

Principled Ways of Finding, Analysing and Planning for Communicative Overhead in Interaction Technology for Fashion Industry

Jason Yang, Sean Rintel, and Stephen Viller

School of ITEE, The University of Queensland, QLD 4072 Australia
{jyang,viller}@itee.uq.edu.au, s.rintel@uq.edu.au

Abstract. In this paper, we identify and analyse the problems associated with communicative overheads of a fashion design and manufacturing workflow. We conduct a multi-stage qualitative study to investigate where the rich multiple channels of communication afford and constrain the workflow during remote collaboration. From this study, we define what communication channels we have in our system. We then use the video data that we had collected through user testing, together with the feedback from the video-assisted stimulated recall interviews, to identify three kinds of communicative overheads in our system: novelty of technology, ongoing constraints and operational problems. We analyse each communicative overhead with examples from our video data, and conclude that there are various overheads, some that participants may overcome through learning and familiarity with the system, some that may not be easily overcome, and some that may not even be unique to a remote collaborative environment, as they also occur in a face-to-face collaborative environment. We believe that offering richer communication channels does not necessarily equate to greater efficiency in the collaborative process, and that designers and developers of collaborative systems need to investigate the effect of communicative overheads before introducing new channels to their system, as these overheads may decrease the efficiency of any collaborative experiences.

Keywords: Coordination, workflow, communicative overhead, channel, remote collaboration, mash-up technologies, qualitative study.

1 Introduction

As communication technology advances, new communication channels emerge. These channels can assist coordinating activities, joint problem solving and sharing artefacts of work, but they also raise the potential for new communicative overheads. Communicative overhead is anything that participants could account for as extra communicative work, such as taking more time explain something, increased incidence of repair, or increased levels of meta-communication. In this paper we explore three sources of communicative overheads: novelty of the technology, ongoing constraints, and operational problems.

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 484–491, 2013.

2 Related Work

The choice of communication channel can be determined by both social presence theory and media richness theory. Short, Williams and Christie first used the term 'social presence' [1] to describe the absence of nonverbal communication which reduced social presence in mediated communication. As computer-mediated communication (CMC) has evolved, the social presence is said to become a way for the individuals to represent themselves and exchange messages, and how those messages are interpreted by others in an online environment [2]. Media richness theory suggests that the effectiveness of the communication is dependent on the selection of an appropriate medium [3], based on the degree of ambiguity users are willing to accept in any given communication situation [4]. It is a channel-based approach to the selection of an appropriate technology for a given situation.

Burgoon proposed the principle of interactivity framework that offers understanding of the interrelationships of verbal and nonverbal cues within CMC [5]. The principle of interactivity assumes that the degree of interdependent, contingent, participative, and synchronous interaction provided by a communication interface or experienced by interlocutors, or both, will affect the social judgments and task performance. [5] [6].

According to the 3C collaboration model (communication, coordination and cooperation) [7], individuals need to communicate and negotiate with others in order to make decisions and exchange information. They also need to minimise conflict to prevent loss of communication while they are coordinating and organising themselves to work together. "The need for renegotiating and for making decisions about unexpected situations that appear during cooperation may demand a new round of communication, which will require coordination to reorganize the tasks to be executed during cooperation." [8]. Neale and Carroll interpreted that coordination can sometimes be viewed as overhead or operating cost while completing interactive group activities [3].

The challenge of supporting effective collaborative group work in a distributed context depends on the level of communicative overheads that are embedded within each communication channel. Communicative overheads have an impact upon interactivity, especially when they are seen as getting in the way of the task. That being said, it is overly simplistic to equate communicative overheads with a negatively valanced sense of trouble or that they are necessarily show-stoppers to adoption [9].

In this paper, we are focused on finding, analysing and planning for communicative overheads within a Textual Visual Tangible Multi-touch (TVTM) system [10] in a remote collaboration setting. The TVTM system consists of two networked multi-touch tables that present users with a shared common workspace. Each table offers an array of simple design editing tools, a basic pen tool for writing and annotating, and a live webcam for demonstrating physical artefacts.

3 Previous Study

We previously identified different modes of interaction, such as the manipulation of physical artefacts, as opposed to textual, visual or tangible encoded virtual representation of these artefacts. We also identified collaboration issues in the textile industry through interviews and observational studies, in terms of classification of context in a fashion design environment [11].

Based on our pilot studies of collaboration, and the use of currently available technologies such as natural user interface (NUI) technology [12] and tangible user interface (TUI) technology [13] [14] in a fashion design environment, we have developed the TVTM system; a 'mash-up' system from a selection of currently available technologies that will allow us to utilise the multi-touch and tangible aspects of the technology, which is more effective than a single standalone technology such as video conferencing [15].

We have also conducted an experiment that required the solving of a design problem using a traditional fax machine versus our TVTM system between two remote locations. The results showed our TVTM system was significantly more efficient, and minimised communicative overheads.

4 Qualitative Study

We have conducted a multi-stage qualitative study to investigate and evaluate the use of the TVTM system and hardware for supporting remote interaction and collaboration in the fashion industry, and also to find and evaluate communicative overheads within communication channels. This study also explored the phenomenon of the gesture-based interaction around the multi-touch gesture technology. For our first stage, we were evaluating the TVTM software with pairs of design students undertaking collaborative tasks. Stage 2 of the study used the video recording from stage 1 to explore and review the interactions in more detail in terms of communicative overheads. Stage 3 of the study involved a game playing process to explore common interaction themes in the video data.

4.1 Methods and Settings

We conducted our user testing with a group of female students in their second and final years of a fashion design course at a local TAFE collage. There were eight participants in total and each user testing session was conducted in pairs, giving four groups in total.

The first stage of the study was to evaluate the participants undertaking a design critique collaborative task (which is a part of their course assessment). Each participant was assigned a TVTM system to complete the task. The participants were shown a short video clip detailing the functionality of the interface, and operating instructions prior to each user testing session.

The participants were encouraged to "think out loud" during the entire session, and the interaction between the participants was studied and captured by recording video during each testing session. At the conclusion of the first stage of the study, we reviewed our observation notes together with the recorded footage to create targeted sets of questions for the video-assisted stimulated recall interviews. This style of interview technique uses video footage to assist participants to recall their experience of user testing during post-event interviews.

We began conducting the second stage one week after we had completed our stage 1 study. We requested our participants come back for the interviews in the same pairs as in stage 1 of the study. The entire interview procedure was described to the participants, followed by their viewing a video clip of their specific user testing session.

While the participants were watching video footage of their participation in the first stage, we began a series of questions directly related to some of the interactions that were observed during their user testing session. The majority of these questions were related to why there were certain interactions with the system or with their partner during the user testing.

We asked questions related to the participants' thought processes at a specific moment during the user testing, to obtain a clear understanding of their design (thought) process while they were critiquing each other's work. We also asked the participants to describe their experience during the entire collaborative process, and obtained feedback regarding the usability of the TVTM system.

4.2 Defining Communication Channels

Our TVTM system offers a mash-up of multiple communication channels. These channels include an audio channel, a shared screen visual channel, a live webcam video feed channel, a multi-touch & gesture enabled visual interface channel, a virtual keyboard input channel and a hand writing/drawing input channel.

4.3 Results – Discovering Overheads

Through observation and analysis of the video data that was obtained, it was determined that there were three different kinds of overheads. They are the novelty of the technology, ongoing constraints, and operational problems. We will discuss each kind of overhead, and relate it to some specific events that were observed in the video data and the participants' responses from the video assisted stimulated recall interviews.

Novelty of the Technology

Example 1: Participant A was trying to convey to participant B a fabric image she wanted for her alternative swimwear design. Participant A was using her finger to tap on the image thinking it would bring up the image that she selected (Fig. 1), participant A said *"I don't think it is working"*. In the meantime, participant B did not know what was happening, so participant B asked participant A *"what is going on? what are you trying to do?"*.

It appeared as though Participant A was trying to select the image using the traditional 'point and click' method with which she was familiar. During this period, Participant B was unaware of the problem that Participant A was experiencing. The interviewer then encouraged Participant A to think out loud so that Participant B could be aware of the situation.

Participant A: *"I realised that I needed to drag the images out of the image selection panel in order to bring up the fabric image. I then explained my thought process to participant B while dragging the image out to the work area."*

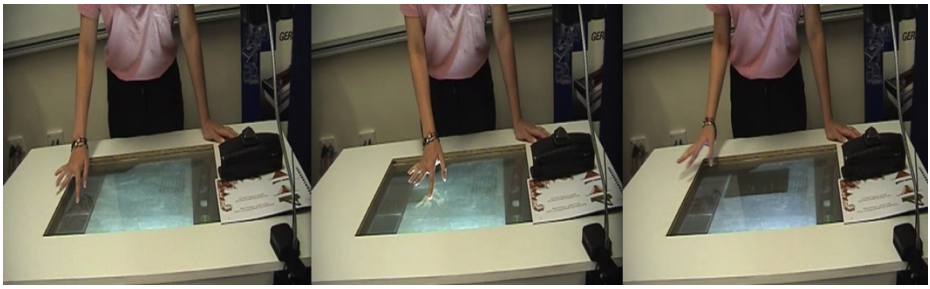


Fig. 1. Participants tapped on the screen first, but nothing happened, a few seconds later she realised she needed to drag the image out

Example 2: Participant C appeared to be exploring the functionality of the TVTM system as she was repeatedly tapping and browsing on the screen, working through the available tools to select the most appropriate.

The interviewer asked Participant C to confirm what she was trying to do at that time.

Participant C: *"I did not know which tools I needed so I was trying to find another way to show participant D my alternative design so she can understand it"*

In this particular example, participant C was unable to recall what design tools the TVTM system offered. Participant C had to click on each design tool in turn, to remember the function of each tool, so as to identify the best tool to illustrate the alternative design to participant D. While participants are exploring and trying to familiarise and learn the system, it causes communication delay/overhead.

Ongoing Constraints

Participant E was browsing through some images, she repeatedly browsing up and down using her finger on the screen.

Participant E: *"Can you see what I am doing?"*
Participant F: *"Yes I can but I have no idea what you are doing though."*
Participant E continued to browse the images.
Participant F: *"What are you doing (Participant E's name!)"*

In this example, participant F was watching participant E's interaction with the TVTM system through the shared screen ability of the system. Participant F did not know what participant E was trying to do, since participant E did not initially

verbalise her thought process. There are many examples in our video data where similar situations occurred. This is due to the fact that if one of the participants worked 'individually' and 'silently' (for example searching an image) and did not "think out loud", then the other participant had no clue about what is happening. This particular type of overhead is not unique to this environment using the TVTM system however, it could also occur with participants in a face-to-face collaborative environment.

Operational Problems (Hardware & software)

Example: Participant I was showing an actual fabric sample over the webcam to participant J. The live webcam video feeds were transmitted over to the system. Participant J was having trouble seeing the live video feed (Fig 2) so she said, *"could you lift it up a bit?"*

Fig. 2. Participant I (left) raised the fabric samples up towards the webcam

The interviewer then asked Participant I to clarify the nature of the problem she was experiencing.

Participant J: *"the image on the screen is too small, so I asked participant I to lift it up so the camera can zoom in a bit"*

In this example, there was a hardware and software limitation, as the resolution of the webcam did not provide sufficient quality of live video feed. Participant J requested participant I to lift the fabric closer to the webcam, so the image would appear larger, with more fine detail visible. This lost time represents an overhead during the collaboration.

Novelty can potentially be overcome through training, experience and time. Ongoing constraints and operational problems, which can be either hardware or software related, can also introduce additional overheads.

5 Discussion

From our previous observational study on a clothing factory, we noticed clothing designers were still using both telephone and facsimile to communicate with their sub-contracted manufacturers in China.

Communication by telephone can be classified as 'mediated interaction' [5]. It is leaner than face-to-face communication as it removes all the visual information while maintaining verbal and audio cues to communication.

Facsimile can be classified as a written channel; for example, clothing designers were able to send out a design alteration specification quickly over to the overseas manufacturer, especially when there is a time difference between the two remote locations. However, the facsimile communication channel produces overheads. As we showed in our previous study, it was not efficient in a situation like collaborating on a problem-solving task in two remote locations.

The communicative overhead inherent with a facsimile communication channel is primarily due to asynchronous communication. Facsimile also lacks important properties, such as the ability to manipulate text and graphic images [16]. From our interview with the clothing designers, they usually received a facsimile response from the remote manufacturer within normal business hours. Sometimes they sent and received facsimiles that contained the same graphic images that had previously been facsimiled, along with additional hand written notes, or annotations around the images on the original facsimile.

In this study, we have attempted to identify the communicative overheads associated with the deployment of a mash-up of rich multiple channels within our TVTM system. This mash-up of multiple channels provides both affordance and constraint to the workflow. In this context, the richer communication channels afford greater efficiency and interactions for remote collaboration, whereas the communicative overhead is the constraint.

In real world settings, people who engage in collaboration apply various channels of interactions to cope with these complexities of workflow; in most cases, overhead or operating cost is involved [1, 3]. If the group collaboration suffers from a high level of overheads, it is possible that the time lost due to these overheads may not be regained by the use of the richer communication channels, leading to a potentially less efficient collaborative experience, and possible abandonment of the joint activities.

Designers and developers need to consider that the notion of overheads is related to the notion of cognitive overload, which can be identified by tracking participants' thought processes while they are completing a task.

New channels may not necessarily create overheads or eliminate existing overheads. From the results of our study, we identified an overhead not directly attributable to the use of any specific channel, which may equally occur during face-to-face or remote collaboration.

Designers and developers of CMC systems can bring new channels into use for particular tasks, however overheads have an influence on how effective new channels are for users working remotely. Therefore they should not necessarily expect an immediate improvement to workflow.

We conclude that our study has shown that offering richer communication channels does not necessarily equate to greater efficiency in the collaborative process.

References

1. Short, J., Williams, E., Christie, B.: The social psychology of telecommunications, 195p. Wiley, London (1976)
2. Lowenthal, P.R.: The evolution and influence of social presence theory on online learning. Online education and adult learning: New frontiers for teaching practices. IGI Global, Hershey (2009)
3. Neale, D.C., Carroll, J.M., Rosson, M.B.: Evaluating computer-supported cooperative work: models and frameworks. In: Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work. ACM (2004)
4. Daft, R.L., Lengel, R.H.: Organizational information requirements, media richness and structural design. Management Science 32(5), 554–571 (1986)
5. Burgoon, J.K., Bonito, J.A., Ramirez, A., Dunbar, N.E., Kam, K., Fischer, J.: Testing the interactivity principle: Effects of mediation, propinquity, and verbal and nonverbal modalities in interpersonal interaction. Journal of Communication 52(3), 657–677 (2006)
6. Walther, J.B., Gay, G., Hancock, J.T.: How do communication and technology researchers study the internet? Journal of Communication 55(3), 632–657 (2006)
7. Ellis, C.A., Gibbs, S.J., Rein, G.: Groupware: some issues and experiences. Communications of the ACM 34(1), 39–58 (1991)
8. Fuks, H., Alberto, B.R., Marco, A.G., Lucena, C.J.: Applying the 3C model to groupware development. International Journal of Cooperative Information Systems 14(02n03), 299–328 (2005)
9. Rintel, E.S.: Conversational management of network trouble perturbations in personal videoconferencing. In: Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction. ACM (2010)
10. Yang, J., Muhlberger, R., Viller, S.: TVTM: A case study and analysis of 3 virtual representations to support remote collaboration within the fashion industry. In: 2010, School of Information Technology and Electrical Engineering (ITEE). The University of Queensland,
11. Yang, J., Dekker, A., Muhlberger, R., Viller, S.: Exploring virtual representations of physical artefacts in a multi-touch clothing design collaboration system. In: Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7. ACM (2009)
12. Jacob, R.J., Girouard, A., Hirshfield, L.M., Horn, M.S., Shaer, O., Solovey, E.T., Zigelbaum, J.: Reality-based interaction: a framework for post-WIMP interfaces. In: CHI-Conference 2008. Association for Computing Machinery Inc. (2008)
13. Hinckley, K., Pausch, R., Proffitt, D., Kassell, N.F.: Two-handed virtual manipulation. ACM Transactions on Computer-Human Interaction (TOCHI) 5(3), 260–302 (1998)
14. Patten, J., Ishii, H., Hines, J., Pangaro, G.: Sensetable: a wireless object tracking platform for tangible user interfaces. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM (2001)
15. Yang, J., Viller, S., Rintel, S.: Outsourcing: Mashing up design methods and technologies in the fashion industry. In: Proceedings: PIN-C 2012 (2012)
16. Sproull, L., Kiesler, S.: Connections: New ways of working in the networked organization. MIT press (1992)

Continuing On-Premise or Adopt On-Demand? An Empirical Study of ERP Adoption in SMEs

Fan Zhao and Elias Kirche

Florida Gulf Coast University, FL, USA
{fzhao,ekirche}@fgcu.edu

Abstract. Vendors are now expanding into the small to mid-size organizational markets with simplified, less-risk, less-reward systems However, SMEs, who adopted ERP on-premise solutions before, are now facing a dilemma: continuing with ERP on-premise upgrades or switch to on-demand solutions. This study surveyed Chief Information Officers of SMEs with respect to indicators of ERP adoptions. Cost, reduced demand for own IT resources, outage/accessibility and performance were found to be the most critical and important factors to assess ERP adoptions for SMEs.

Keywords: ERP, Cloud Computing, Adoption, Upgrade, on-demand.

1 Introduction

ERP solutions have evolved from reorder point and materials management focused systems (ROP and MRP) based on mainframes to responsive and integrated resource planning systems based on client-server architectures and web platforms. Through data standardization and process integration, ERP systems have the potential to facilitate communications and co-ordination, enable the centralization of administrative activities, reduce IS maintenance costs and increase the ability to deploy new IS functionality [11]. ERP systems are adopted for strategic, technical, and operational reasons including: providing an integrated enterprise-wide application with real-time data access available across the entire organization, simplifying and standardizing systems and business processes, and replacing legacy systems. When they are well implemented, ERP systems are able to bring operational, managerial, strategic, information technology (IT) infrastructure and operational benefits to their customers [25]. The motivations which have become more dominant include pressure to keep up with competitors on a global scale, mergers and acquisitions and the need for restructuring, and combining and integrating business processes. ERP systems have spread rapidly among organizations. According to Gartner [10], worldwide spending on enterprise application software will total $120.4 billion in 2012, with ERP the largest application software market expected to reach $24.9 billion.

The fast changes in technologies during the last decade dramatically altered the nature of the demand requiring more reliable services, rapid customization and on time delivery of goods and services, forcing many small and medium-sized enterprises (SMEs) to adjust their business strategies and to adopt a better technology to increase

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 492–500, 2013.

the productivity of business processes. For SMEs, the advantages gained by adopting ERP systems are obvious. Gattiker and Goodhue [11] describe the ERP benefits in following aspects: (1) improving the integration of information flow between sub-units; (2) centralizing the administrative activities, such as accounts payable and payroll; (3) reducing costs of system maintenance; (4) increasing the ability to deploy new IS functionality; (5) enabling transformation from inefficient business processes to accepted best practices. With these insights, an ERP system seems to be one logical alternative to be pursued by SMEs [21]. And with the majority of large companies having already implemented ERP systems, today ERP vendors, such as SAP and Oracle, are shifting their focus towards SMEs with scaled-down (and less expensive) versions of ERP solutions. Other important reasons have encouraged the interest of ERP vendors towards SMEs. These include the need for SMEs to be integrated into the supply chain at risk of becoming less competitive, the large market represented by the number of SMEs (90 percent of all businesses in the USA are SMEs, Sledgianowski et al. [26]) compared to the number of large enterprises, and the advancement in technologies with corresponding decrease in hardware costs. Another factor in support of ERP systems is that often information visibility is an important aspect in supply chain performance and the lack of an adequate technology will debilitate the SME's ability to adapt to changes imposed by customers or suppliers with unfavorable impact in its competitiveness. However, the use of commodity software, such as ERP systems, may force a more rigid structure on a SME and thus weaken its flexibility and agility to adapt as a competitive advantage. Therefore, it is critical to SMEs to investigate their performance after the implementation of ERP systems.

In recent years, after many SMEs adopted various ERP packages from different ERP vendors, a new technology model, called Cloud Computing, became possible for sourcing ERP systems with the potential to have a significant impact for SMEs' competitiveness. Cloud computing is an emerging business and technology concept to support an on-demand delivery of computing, storage and applications over the Internet. A recent IDC report shows global revenue in cloud software market reaching $22.9 billion and it will grow to $67.3 billion in 2016 [18]. This projection includes revenue generated by the shift from on-premise to on-demand providers as well as by the planning and architecture behind the shift.

Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of applications and resources and may be offered in one of the following service models [4, 16, 17, 28, & 29]:

- Software as a service (SaaS): providing software subscription services
- Storage as a service: providing remote storage resource services
- Database as a service: providing remotely hosted database services
- Information as a service: providing remotely hosted information services
- Process as a service: providing business processes based on remote resources
- Testing as a service: providing testing services for local or remote systems
- Platform as a service (PaaS): providing a complete platform to support application development, interface development, database development, storage, information and testing
- Infrastructure as a service (IaaS): providing a service to access computing resources remotely

- Security as a service: providing core security services remotely over the Internet
- Integration as a service: providing a complete integration stack service

Specifically for ERP systems, Beaubouef [2] indicated that are three cloud models for adoption:

- Software as a service (SaaS): a subscription model for small customers who share hardware.
- Hosted ERP: a typical solution for large customers who have separate hardware and instances.
- Hybrid ERP: a combination solution that maintains on-premise software as well as integrated a degree of on-demand services.

As high-speed internet has become more accessible to businesses, cloud models have helped businesses with limited IT resources take advantage of technology to improve business processes. Typically, a firm would rent licenses or access the software from an application service provider (ASP) that actually runs on servers or devices owned and maintained by the ASP. These large datacenter facilities with redundant layers of power and data security are often too expensive for SMEs to have under the on-premise model. Also, the ASP is generally responsible for maintaining and updating the software, and often includes some level of support for users in the monthly fee. This model has certain benefits for SMEs that can attenuate some of the problems associated with maintaining purchased software in-house. For example, SaaS allows ASP to maintain their offerings consistently by automating testing, monitoring, maintenance and upgrades without sending out constant updates that need to be applied by end users. Also, SaaS allows smaller firms with limited (or no) IT staff to benefit from the economies of scale and efficiencies implemented by the ASP. Equally important, SaaS allows companies to pay to use the software they need, without making a huge investment in IT infrastructure for servers, software, etc., by "renting" access to what they need and paying monthly, quarterly or annually.

Additionally, according to Gartner's report [1], more and more companies are considering on-demand services in different applications (Figure 1).

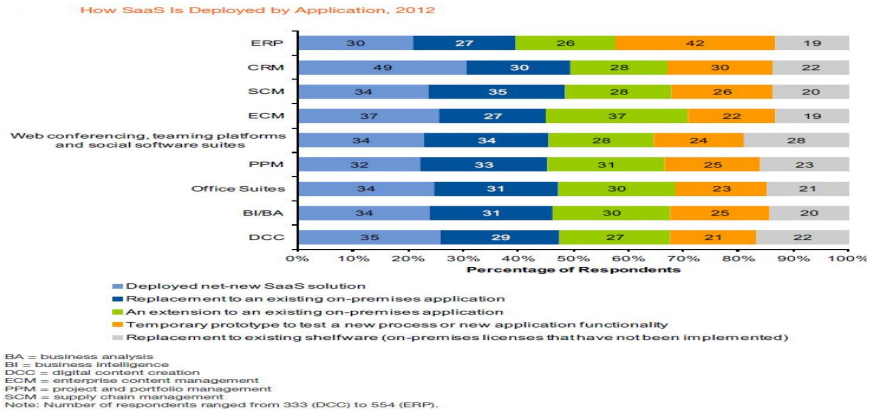


Fig. 1. How SaaS Is Developed by Application

Consequently, SMEs with existing ERP on-premise packages will be facing a new challenge when a new version of the on-premise ERP is released or when they require a substantial IT infrastructure investment. The question for management is whether it will decide in favor of the ERP upgrade or switch to an on-demand ERP. To date, there is no empirical research that has been done to evaluate and compare the two options for SMEs. Therefore, the purpose of this study is to investigate the on-premise and on-demand options for SMEs sourcing ERP solutions and summarize the key factors influencing decision making of the ERP future solutions among CIOs and IT managers.

2 Theoretical Background

The primary purpose of ERP upgrades is to take advantages of new technologies and business strategies to ensure that the organization keeps up with the latest business development trends. Therefore, the decision to upgrade ERP is usually not driven by code deterioration or anticipated reduction in maintenance costs alone. According to an AMR research [27], 55% of upgrades were voluntary business improvements triggered by the need for new functionality, expansion or consolidation of systems; 24% of upgrades were triggered by technology stack changes; 15% of upgrades were forced by de-support of the running version of software to avoid vendor support termination [7]; and 6% of upgrades were triggered by bug fixes or statutory changes. The cost of ERP upgrades is high [20]. Swanton [27] mentioned that the cost of each upgrade includes: 50% of the original software license fee and 20% of the original implementation cost per user, which means over $6 million dollars for a 5,000-user system. Typically, each ERP upgrade requires eight to nine months of effort with a team the equivalent of one full-time employee per 35 business users. The ERP-adopting organization does not have to develop and re-write the ERP system itself but rather it replaces (or upgrades) the old version with a readily available new version from the ERP vendor. However, a lack of experience may cause the costs and length of the upgrade project to approach or even exceed those of the original ERP implementation effort. Every three years, a major ERP upgrade and several small upgrades are typically needed to keep the system running smoothly. Organizations will spend a significant amount of money on each ERP upgrade project, which may place the system out of reach (or postponed) for many SMEs with limited financial resources and shortage of skilled personnel.

However, cloud computing supports an on-demand delivery of ERP over the Internet with competitive implications for SMEs. Typically, ERP functions will be served out of the cloud, and the license fees and/or maintenance fees will be charged based on the use of the software, such as number of users and ERP functions/modules. Companies will no long need to purchase the software and to keep IT staff, which will make the initial investigation of EPR less expensive. This means SMEs now can access the same technology until recently only available to larger firms, and provide the same or superior level of services than larger competitors with the on-premise

ERP classic model. According to Gheorghe and Lupasc [12], cloud ERP solution will save about 30% project time after analyzing the critical path of both classic and cloud ERP solutions. Table 1 summarizes the benefits derived from cloud computing ERP systems.

Table 1. Benefits of On-demand ERP solutions

Benefits of on-demand ERP	Reference
Short implementation cycles	[12]; [13];
Low entry costs	[19]; [13]; [3]
Reduced demand for own IT resources	[13]; [3]
Elasticity/Flexibility of services	[19]; [13]; [23]
Scalability	[19]; [10]
Focus on core business	[13]

However, there are also several challenging issues for the adopters of on-demand ERP, and cost is always an important one. Although companies can adopt on-demand ERP solutions at a much lower initial cost and lower total cost of ownership comparing with the traditional ERP solutions, the on-going monthly or annual payment will result in a higher budget cost because of the subscription payment method [22]. Therefore, the subscription method is of critical importance for this technology model to be sucessfuly adopted and widely accepted by SMEs. Table 2 summarizes the challenging issues derived from cloud computing ERP systems. As more companies who adopted on-premise ERP solutions are considering on-demand solutions, a research study is needed to give guidelines to SMEs when they are faced with this dilemma.

Table 2. Challenging Issues of On-demand ERP solutions

Challenging Issues of on-demand ERP	Reference
Outage/Accessibility	[15]; [22]
Security	[15]
Performance	[15]
Integration	[15]; [22]
Cost	[15]; [22]
Control of the system	[22]
Flexibility of customization	[22]

This study seeks to provide a comprehensive understanding of two ERP solutions: on-premise and on-demand, collecting CIOs' perspectives in different SMEs according to their experiences from their organizations' adoptions of ERP solutions.

3 Research Methodology

The survey questionnaire was sent to CIO/IT managers in 272 SMEs, who are currently using on-premise ERP solutions. Company size is measured by the number of employees [6]. The objective is to conduct an analysis to confirm the underlying structure of the factors involved in the on-demand and on-premise ERP adoption decisions. In this study, an SME is one with fewer than 500 employees. A total 63 survey responses were obtained after removing 7 incomplete questionnaires. The distribution of the respondents was as follows: gender (male: 76%, female: 24%), current EPR systems in use come from many vendors, such as SAP, Microsoft, Syspro, NetSuite, etc. The number of employees range from 36 to 483 (64% small companies and 36% mid-size companies). A wide variety of industries were represented in the responses (shown in Table 3).

Table 3. Industry Summary

Industry Sector	**Number of companies**
Manufacturing	18
Information technology	19
Retail	5
Health Services	16
Insurance	5
Total	**63**

To substantiate our measurement model for the factors identified from the literature (Tables 1 and 2), content validity was evaluated by a group of experts. Six experts from two universities and four organizations were asked to read the survey and decide whether in their opinion it measured what its name suggested. The survey questionnaires were improved according to comments from the experts. The final survey was returned to the experts and approved by all. The coefficient alpha [8] to measure the internal consistency reliability of the survey instrument ranged from 0.73 to 0.85, indicating good reliability according to the acceptable level (0.7) suggested by Bryman [5]. An item-to-total correlation was conducted to evaluate the convergent validity. The literature suggests that a measurement variable may be considered for removal if the item-to-total correlation is less than 0.4. One question was removed because the item-to-total correlation was lower than 0.4. Construct validity is one of the most significant scientific advances of modern measurement theory and practice [14]. It measures how well the instrument measures the concept of interest. Confirmatory factor analysis (CFA) was conducted to examine construct validity. All but one construct that emerged from the factor analysis showed high construct validity. The overall results of the analysis confirmed the factors in the ERP adoption solutions presented in the literature. Additionally, the survey asked the respondents to rank the factors in order of importance. The results are explained in the next section.

4 Results and Discussion

Survey results focused on factors involved at ERP adoptions between on-premise and on-demand solutions. For each factor, a brief description and a 5-level rating scale ranging from extremely critical and important in ERP adoption evaluation to neither critical nor important in ERP adoption selections was provided. The numerical anchors for the 5-level rating scale are indicated in Table 4.

Table 4. Survey Averages by Factor

FACTOR	AVERAGE (1 – 5 scale)
Cost	4.93
Reduced demand for own IT resources	4.51
Outage/Accessibility	4.46
Performance	4.07
Control of the system	3.85
Elasticity/Flexibility of services	3.78
Short implementation cycles	3.49
Security	3.24
Scalability	3.03
Integration	2.70
Flexibility of customization	2.46
Focus on core business	2.23

According to the average rating in Table 4, the results indicate that there are top four factors, which is extremely critical and important in ERP adoptions for SMEs, identified by CIOs facing to the options of on-premise or on-demand solutions: cost, reduced demand for own IT resources, outage/accessibility and performance. The second group, which is critical and important, includes five factors: control of the system, elasticity/flexibility of services, short implementation cycles, security and scalability. The rest of the factors, integration, flexibility of customization, and focus on core business, receive low rating from CIOs.

There are several important results summarized from data. First, besides the three factors (cost, reduced demand for own IT resources, and performance) that are commonly regarded as critical success factors in ERP selection, outage/accessibility is rated as one of the highest factors. It is obvious that system outage is closely correlated with on-demand adoptions and may represent a high concern for SMEs with intentions of adopting this ERP solution. In fact, after the outage issue of Microsoft Azure, CIOs realized the importance of system accessiblity for on-demand solutions. Second, security was emphasized by several researchers in their studies of cloud computing [15, 24]. However, according to our data, CIOs only rated security in the second importance group. Third, most of the small companies prefer adopting on-demand solutions while mid-size companies are more hesitante to switch from current system to on-demand ERP.

The results of this study provides SMEs with valuable knowledge that might prompt them to make significant changes in their current ERP-related processes and the acquisition of enterprise packaged software, which could result in substantial benefits. The findings have important theoretical and practical implications. Our understanding of the ERP related issues in SMEs helps both researchers and practitioners understand how ERP system adoption and implementation impact the competitiveness of SMEs. To researchers, this study shows the relationship between on-premise and on-demand ERP adoptions and give the possibility in future research to expand the model. To practitioners, this study shows some critical issues of ERP adoption for SMEs. It allows CIOs and other IT managers to think thoroughly before they make decisions in ERP adoptions. Additionally, guidelines are provided to IT managers regarding the ERP adoption issue between on-premise and on-demand models. Furthermore, the ERP vendors will benefit from this study in developing better ERP solutions tailored to SMEs specific requirements, and improve the cost structure of their cloud models for higher adoption rates among SMEs.

References

1. Anderson, E., Britz, B., Bell, W.: User Survey Analysis: Using Cloud Services for Mission-Critical Applications (September 28, 2012), http://www.gartner.com/id=2182216 (retrieved February 3, 2013)
2. Beaubouef, B.: Cloud ERP – New Dog, Same Fleas (May 22, 2011), http://gbeaubouef.wordpress.com/2011/05/22/cloud-erp/ (retrieved February 5, 2013)
3. Bhardwaj, S., Jain, L., Jain, S.: Cloud Computing: A Study of Infrastructure as a Service (IaaS). International Journal of Engineering and Information Technology 2(1), 60–63 (2010)
4. Branzei, O., Vertinsky, I.: Strategic pathways to product innovation capabilities in SMEs. Journal of Business Venturing 21, 75–105 (2006)
5. Bryman, A.: Social Research Methods. Oxford University Press, New York (2004)
6. Chow, S., Holden, R.: Toward an understanding of loyalty: The moderating role of trust. Journal of Managerial Issues 9(3), 275–298 (1997)
7. Craig, R.: Laurier enterprise system upgrade. In: Proceedings of International Conference of Information Systems, Charlotte, pp. 654–662 (1999)
8. Cronbach, L.: Coefficient alpha and the internal structure of tests. Psychometrika 16, 297–334 (1951)
9. Furht, B., Escalante, A.: Handbook of Cloud Computing. Springer Science + Business Media, New York (2010)
10. Gartner, Inc. Stanford, Connecticut, USA. Press release: Global IT Spending Forecast (June 20, 2012), http://www.gartner.com/newsroom/id/2055215 (retrieved)
11. Gattiker, T., Goodhue, D.: Understanding the plant level costs and benefits of ERP: will the ugly duckling always turn into a swan. In: Proceedings of the 33rd Hawaii International Conference on System Science, Hawaii (2000)
12. Gheorghe, G., Lupasc, I.: Crises Solutions for SMEs Case Study: ERP Cloud Vs. Classic Solution. In: International Conference "Risk in Contemporary Economy" Galati, Romania, SIIIth edn., pp. 149–156 (2012)

13. Karabek, M., Kleinert, J., Pohl, A.: Cloud Services for SMEs - Evolution or Revolution? Business + Innovation, 26–33 (2011)
14. Kerlinger, F.: Foundations of Behavioral Research. Rinehart and Winson, Fort Worth (1986)
15. Kim, W.: Cloud Computing: Today and Tomorrow. Journal of Object Technology (January/February 2009)
16. Kim, W., Kim, S., Lee, E., Lee, S.: Adoption Issues for Cloud Computing. In: Proceedings of the 7th International Conference on Advances in Mobile Computing and Multimedia, pp. 2–5. ACM, New York (2009)
17. Linthicum, D.: Cloud Computing and SOA Convergence in Your Enterprise. Addison Wesley, Boston (2010)
18. Mahowald, R., Sullivan, C.: Worldwide Saas and Cloud Software 2012-2016 Forecast and 2011 Vendor Shares. International Data Corporation (IDC) (2012), http://www.idc.com/getdoc.jsp?containerId=236184#.USTXoh3kseB (retrieved)
19. Marston, S., Li, Z., Bandyopadhyay, S., Zhang, J., Ghalsasi, A.: Cloud computing - The business perspective. Decision Support Systems 51(1), 176–189 (2011)
20. Montgomery, N.: Build your business case for upgrades by adding functionality. Computer Weekly, 16 (2004)
21. Olson, K., Satre, P.: ERP for SMEs – is proprietary software an alternative? Business Process Management 13(3), 379–389 (2007)
22. Panorama Consulting Group.: SaaS ERP vs. on-premise ERP software: Six key differentiators (March 12, 2009), http://searchcloudapplications.techtarget.com/news/1350684/SaaS-ERP-vs-on-premise-ERP-software-Six-key-differentiators (retrieved February 20, 2013)
23. Rittinghouse, J., Ransome, J.: Cloud Computing: Implementation, Management, and Security. CRC Press (2009)
24. Saeed, I., Juell-Skielse, G., Uppstrom, E.: Cloud Enterprise Resource Planning Adotions: Motives & Barriers. In: International Conference on Research and Practical Issues of Enterprise Information Systems (2011)
25. Shang, S., Seddon, P.: A comprehensive framework for classifying the benefits of ERP systems. In: Proceedings of the Americas Conference on Information Systems, Long Beach, CA (2000)
26. Sledgianowski, D., Tafti, M., Kierstead, J.: SME ERP system sourcing strategies: a case study. Industrial Management and Data Systems 108(4), 421–436 (2008)
27. Swanton, B.: Build ERP upgrade costs into the business change program – not the IT budget. Computer Weekly, 28–28 (September 21, 2004)
28. Velte, A., Velte, T., Elsenpeter, R.: Cloud Computing, A Practical Approach. McGraw-Hill (2010)
29. Wang, L., von Laszewski, G., Younge, A., He, X., Kunze, M., Tao, J., Fu, C.: Cloud Computing: A Perspective Study. New Gneration Computing 28, 137–146 (2010)

Erratum: Exploring Children's Attitudes towards Static and Moving Humanoid Robots

Fang-Wu Tung[1] and Tsen-Yao Chang[2]

[1] Department of Industrial and Commercial Design, National Taiwan University of Science and Technology, 43,Sec. 4, Keelung Rd., Taipei, 106, Taiwan
[2] Department of Creative Design, National Yunlin University of Science and Technology, Yunlin County, 64002, Taiwan
fwtung@mail.ntust.edu.tw, changty8908@gmail.com

M. Kurosu (Ed.): Human-Computer Interaction, Part III, HCII 2013, LNCS 8006, pp. 237–245, 2013.

DOI 10.1007/978-3-642-39265-8_56

The originally published version of Table 1 contains an error. This erratum shows the corrected version of Table 1.

Table 1. The experimental stimuli

Image / Behavior	R1	R2	R3	R4	R5	R6	R7	R8	R9
									
Speech					v		v		
Gesture	v		v	v		v	v	v	v
Facial expression					v		v		
Locomotion	v	v	v	v		v			v

The original online version for this chapter can be found at
http://dx.doi.org/10.1007/978-3-642-39265-8_26

Author Index

Abdelnour-Nocera, José III-146
Abed, Mourad V-141
Abou Khaled, Omar IV-157, IV-167, V-19
Acharya, Subrata II-26, III-3
Ackermann, Daniel III-446
Ahn, Jong-gil II-249, III-87
Akihiro, Ogino V-372
Akiyoshi, Masanori I-159, II-444
Alexandris, Christina IV-3, IV-13
Alfredson, Jens I-221
Almutairi, Badr IV-23
Altaboli, Ahamed I-549
Alves, Aline Silva III-324
Alves, Paulo Henrique Cardoso II-3
Amandi, Analía A. III-107
Amer, Hossam III-352
Amin, Rahul III-97
An, Kwang-Ok V-3
Anastasiou, Dimitra IV-32
Andujar, Marvin II-335
Angelini, Leonardo II-531, V-19
Angulo, Julio III-10
Ann, Sangmin V-441
Arantes, Ana Cláudia Costa II-370
Ariya, Kanako I-275
Armentano, Marcelo G. III-107
Arning, Katrin III-49
Arnold, Todd I-593
Arya, Ali IV-215
Asaba, Nobutake I-565
Atia, Ayman II-561
Attard, Judie V-122

Bach, Cédric I-521
Badioze Zaman, Halimah IV-523
Baguma, Rehema III-249
Bahn, Sangwoo IV-561, IV-594, IV-610, IV-618
Baranak, Andrew III-285
Barbé, Jérôme IV-429
Barbosa, Simone Diniz Junqueira I-3, I-460, IV-439
Barnett, Julie I-166
Barroso, João V-39
Beldad, Ardion III-371
Bell, Carrie III-285
Bellini, Pierfrancesco III-259
Belluati, Maurizio III-277
Benetti, Angelo II-428
Benford, Steve II-376
Bengler, Klaus II-578, II-596
Bergmann, Francine B. III-117
Bergmans, Anne II-541
Bernhaupt, Regina I-521
Bevan, Nigel I-281
Beyerer, Jürgen IV-252
Bilgen, Semih I-310
Billman, Dorrit I-193
Bim, Sílvia Amélia I-3
Blezinger, David V-10
Böck, Ronald V-301, V-381
Borsci, Simone I-166, I-203
Borum, Nanna III-418
Boscarioli, Clodis I-3
Bourimi, Mohamed III-39
Boy, Guy Andre IV-629
Brat, Guillaume I-290
Braun, Andreas IV-147
Braun, Roman III-446
Brejcha, Jan I-13
Bröhl, Christina III-127
Brooks, Anthony L. III-418
Brouwers, Aurélie III-136
Brunk, Sören V-46
Brunner, Simon IV-388
Bruno, Ivan III-259
Burghardt, Manuel I-176
Burzacca, Paolo I-241
Bützler, Jennifer III-127

Cagiltay, Nergiz E. I-310
Câmara, Marcela IV-439
Campos, José Creissac I-421
Canedo, Arquimedes I-350
Cantú, Andrea II-316
Caon, Maurizio II-531, IV-167, V-19

Cardozo, Aline Teodosio dos Santos II-370
Carrino, Francesco II-531, IV-157
Carrino, Stefano II-531, IV-167
Caruso, Mario IV-637
Catarci, Tiziana IV-637
Cederholm, Henrik V-403
Cerrato-Pargman, Teresa II-464
Chakraborty, Joyram II-13
Chalon, René V-29
Chamberlain, Paul I-22
Chan, Susy III-183
Chang, Hsien-Tsung III-359, IV-177
Chang, Hsin-Chang II-606, II-654
Chang, Tsen-Yao III-237
Charoenpit, Saromporn II-343
Chen, C.L. Philip V-236
Chen, Huan-Ting IV-177
Chen, Po Chun V-132
Chen, Wei-Ju II-20
Chen, Weiqin IV-186
Chen, Xiuli I-193
Cheng, Cheng IV-243
Cheng, Mingkai IV-243
Cheng, Yun-Maw II-20
Cheung, Yiu-ming V-246
Chien, Tsung-Tien IV-177
Ching, Joanna Sin Sie II-181
Chizari, Mahmoud V-262
Cho, HyunKyoung I-32
Cho, Yongjoo V-339
Choh, Ikuro IV-348
Choi, Heesun IV-618
Christensen, Line G. III-418
Chu, Chi Nung II-386
Chu, Kimberly IV-750
Chuang, Ming-Chuen I-540
Chung, Byung Do III-402
Chung, Ching-Wen III-199
Chung, Donghun II-258
Cincotti, Febo IV-637
Cirilo, Elder José Reioli II-3
Clamann, Michael IV-551
C. Mac, Khoi-Nguyen IV-67
Coats, Brian II-26
Coleti, Thiago Adriano I-184
Corbett, Brendan IV-561
Craig, Claire I-22
Craven, Michael P. II-36, II-189
Crebolder, Jacquelyn I-109
Crowe, John II-36, II-189
Cruciol, Leonardo L.B.V. IV-648
Cusseau, Guillaume II-46

Dahlbäck, Nils II-586
Das, Amitava V-310
Davaasuren, Enkhbat IV-196
David, Bertrand V-29
de Abreu Braz, Priscilla Fonseca IV-439
de Araújo, Fábio Rodrigo Lopes II-3
Debattista, Jeremy V-122
de Carvalho, Alfredo Veiga II-3
Dehlinger, Josh II-66
Delomier, Florent V-29
de Lucena, Carlos Alberto Pereira II-127, II-144, II-370
de Lucena, Carlos José Pereira II-3, II-370
De Marsico, Maria II-351
Dempster, Peter III-269
Deng, Xiaoming V-227
de Oliveira, Helvecio Siqueira II-428
de Oliveira, João Batista S. III-117
de Oliveira, Káthia Marçal I-211
DePalo, Philip II-56
de Souza, Clarisse Sieckenius I-320
Dey, Anind K. V-92
Dharma, Anak Agung Gede IV-568, IV-578
Di Bitonto, Pierpaolo II-484
Dicke, Christina II-551
Di Giovanni, Pasquale III-342
Dixon, Jeremy II-66
Dixon, Shannan DeLany II-66
Dörner, Ralf I-381
Drury, Jill L. IV-658
Duan, Fuqing V-206, V-216, V-280
Duong, Anh-Duc IV-513

Ebert, Achim I-371
Ebisawa, Yoshinobu IV-205
Edirisinghe, Chamari IV-601
Ehrhart, Stefan I-371
Eibl, Maximilian IV-706, V-196
Ekandem, Josh I. II-335
Engel, Jürgen I-300, I-401
Eriksson, Jeanette V-403
Erturan, Yusuf Nasuh I-310
Esat, Ibrahim II-76, V-262
Eshraghi, Mona II-76

Eshraghi, Saba II-76, V-262
Eskildsen, Søren II-361
Espinoza, Matías II-299
Etemad, S. Ali IV-215
Euman, Rob III-146

Fain, Brad III-285
Fallman, Daniel I-128
Fang, Xiaowen II-150, III-183
Farhadi-Niaki, Farzin IV-215
Faria, Jose V-39
Feary, Michael I-193
Federci, Stefano I-203
Fekry, Mohamed II-561
Feng, Shih-Yi IV-177
Fernandes, Hugo V-39
Ferreira, Juliana Jansen I-320
Ferreira, Sérgio IV-678
Fischer, Arthur IV-147
Fleischmann, Albert I-330, III-456
Forbrig, Peter I-51, I-300, I-340
Förster, Ulrich III-446
Frajhof, Leonardo II-127, II-370
Franke, Thomas II-612
Fritsch, Lothar III-68
Froehlich, Elaine IV-668
Fuchs, Christian IV-678
Fuglerud, Kristin Skeide I-41
Fukumoto, Kiyotaka IV-205
Funke, Alexandra V-46
Furuhashi, Takeshi IV-740
Furukawa, Hiroshi III-156

Gabbard, Joseph L. III-217
Gabillon, Yoann I-211
Gadahad, Pallavi Rao II-181
Gailliot, Ava IV-135
Galatas, Georgios IV-43
Galindo, Michel I-51
Gambäck, Björn V-310
Gao, Dekun IV-225
Gao, Yue V-153
Garcia, Franco Eusébio II-229
Germanakos, Panagiotis II-418
Ghedira, Khaled V-141
Ghinea, Gheorghita V-56
Gilbert, Juan E. II-335, III-97
Giménez, Rafael I-256
Go, Kentaro I-77, I-119, I-137, I-500, I-531, IV-456
Göbel, Matthias I-61
Gomez, Randy IV-408
Gonçalves, Gil IV-678
Gong, Yang II-94
Gopalakrishnakone, Ponnampalam IV-601
Goto, Yukio IV-722
Gotoda, Naka II-84
Gotoh, Takuto II-216
Gotoh, Tomomi IV-235
Greene, Kristen K. IV-449
Grinsell, Jon II-46
Gris, Iván IV-97
Grønli, Tor-Morten V-56
Gruss, Sascha V-474
Grüter, Barbara II-323
Gu, Zhenyu I-555
Guercio, Elena III-277
Guo, Ping V-236
Guo, Weimin IV-243

Hachimura, Kozaburo I-611
Hamdy, Aya II-561
Hammer, Jan Hendrik IV-252
Han, Chia Yung IV-291
Handschuh, Siegfried V-122
Hansen, Jarle V-56
Harata, Miho V-319
Harley, Linda III-285
Harris, Helen II-571
Harun, Afdallyna II-376
Hasegawa, Koyo II-647
Hashiyama, Tomonori I-480
Hashizume, Ayako I-68
Hattori, Masatsugu IV-462
Haubner, Nadia I-381
Hayakawa, Seiji I-77, I-119, I-137, I-500, I-531
Hayami, Takehito IV-312, IV-320, IV-330
Hayashi, Kazuya IV-320
Hayashi, Yugo III-20
He, Ning V-153, V-216
Heikkinen, Kari II-239, V-431
Heinroth, Tobias III-59
Helmbrecht, Magnus II-578
Henschen, Lawrence J. II-107, IV-688
Herberhold, Carlotta III-381
Herdin, Christian I-300, I-401
Hermann, Fabian I-256, III-29, V-122

Hermann, Sven I-350
Hernández Rubio, Erika IV-301
Heupel, Marcel III-39
Hijikata, Yoshinori III-314
Hilborn, Olle V-403
Himmel, Simon III-49
Hiramatsu, Yuko II-398
Hirata, Ichiro I-361
Hishina, Masateru I-565
Ho, Yi-Lun III-359
Hölscher, Christoph V-10
Hori, Maiya V-391
Horne, Sara II-76
Hörold, Stephan I-85, I-391
Howes, Andrew I-193
Hrabal, David V-474
Hsieh, Min-Chih III-166
Hua, Lei II-94
Huang, Chi-Yu III-199
Huang, Lixiao IV-561
Huang, Xinpeng I-480
Huang, Yi-Pai IV-506
Huang, Yu Ting II-386
Hufgard, Andreas III-438
Hui, Mingqi V-498
Humayoun, Shah Rukh I-371
Hung, Jian-Yung II-606, II-654
Hung, Matthew III-285
Huseyinov, Ilham N. II-391
Hwang, Sheue-Ling II-606, II-654, III-166, III-193
Hwang, Wonil IV-587
Hwangbo, Hwan I-103
Hyden, Douglas IV-135

Ichino, Junko I-480
Ifukube, Tohru II-398
Iida, Yusuke III-466
Ingold, Rolf IV-157, IV-167
Inoue, Satoshi II-647
Ioannou, Nikolaos IV-358
Ishihara, Makio V-163
Ishihara, Manabu II-474
Ishihara, Yukio V-163
Itakura, Naoaki IV-225
Iteya, Satoru V-411
Ito, Atsushi II-398
Ito, Kyoko V-421
Itou, Junko III-174
Iwai, Yoshio V-270, V-391
Iwata, Mitsuru I-480
Iwata, Naoko I-146
Izumi, Tomoko III-295, V-329

Jackson, France III-97
Jakus, Grega II-551
Jander, Hans I-221
Jeng, Tay Sheng V-132
Jeon, Myounghoon IV-49
Jeon, Wooram IV-594
Jepsen, Henrik W. III-418
Jerg-Bretzke, Lucia V-474
Ji, Yong Gu I-103
Jiang, Yonghan IV-243
Jin, Lianwen V-254
Jochems, Nicole III-127
Johnson, Chris II-20
Johnson, Michael D. II-408
Jokela, Timo II-101
Jokinen, Kristiina IV-32, IV-262
Jonsson, Ing-Marie II-586
Joo, Jaekoo V-62
José, Rui I-421
Ju, Hehua V-171
Jung, Tzyy-Ping V-448

Kaber, David B. IV-551
Kaji, Ken V-372
Kakusho, Koh IV-340
Kalwar, Santosh Kumar II-239, V-431
Kambayashi, Yasushi IV-235
Kamei, Katsuyuki IV-722
Kampf, Constance III-388
Kanayama, Naofumi I-480
Kanenishi, Kazuhide II-503
Karunanayaka, Kasun IV-601
Kashima, Tomoko II-444
Katagami, Daisuke IV-340
Kato, Toshikazu III-466, V-411
Kawabe, Akhiro III-295
Kawahara, Tatsuya IV-408
Kawakami, Hiroshi III-301
Kawakatsu, Hikaru IV-491
Kellen, Vince III-183
Kesdoğan, Doğan III-39
Khaled, Omar Abou II-531, V-186
Khong, Chee Weng IV-750
Kida, Takahiro IV-235
Kim, Da-Hey V-3
Kim, Dongkeun V-339, V-441

Kim, Dongsoo IV-587
Kim, Eun Yi V-362
Kim, Gerard Jounghyun II-249, III-87
Kim, Hyo Chang I-103
Kim, Jongbae V-3
Kim, Jonghwa V-441
Kim, Minyoung V-339
Kim, Youngwon II-249
Kinoshita, Yuichiro IV-456
Kirakowski, Jurek IV-87
Kirche, Elias III-492
Klein, Andreas III-396
Klein, Gary L. IV-658
Kleindienst, Jan IV-59
Kline, Keith III-285
Knecht, Christian III-29
Ko, Deajune V-346
Kobayashi, Noriyuki III-307
Kobayashi, Tomoyuki IV-378
Koh, Yoon Jeon III-402
Kohls, Niko II-163
Kohno, Izumi I-146
Kojima, Kazuaki IV-398
Kojima, Takatsugu IV-340
Komatsu, Tsuyoshi I-585
Kosnik, David E. II-107
Kouroupetroglou, Georgios IV-358
Kraft, Patricia III-411
Krause, Michael II-596
Krems, Josef F. II-612
Kristensen, Kasper III-418
Kroes, Lies II-117
Krömker, Heidi I-85, I-391
Kryssanov, Victor III-20
Ku, Pei-Ying II-606
Kühn, Romina IV-698, V-46
Kümmerling, Moritz I-411
Kunc, Ladislav IV-59
Kuno, Yuki V-72
Kuo, Chih-Chung II-606, II-654
Kuramochi, Toshiya III-314
Kuribara, Takuro IV-416
Kurosawa, Toshifumi IV-272, IV-469
Kurosu, Masaaki I-68, I-95

Labský, Martin IV-59
Laine, Juha II-101
Lalanne, Denis II-531, IV-388, V-186
Lam, Jacob III-418
Laterza, Maria II-484
Le, Hoang-An IV-67
Leal, Jeremy I-109
Leal Ferreira, Simone Bacellar III-324
Lee, Eui Chul V-346
Lee, Hyunji II-258
Lee, Jaehong IV-368
Lee, Ja Young IV-610
Lee, Ju-Hwan IV-49
Lee, Julia C. IV-688
Lee, Unseok IV-281
Lehmann, Simon I-381
Lekkas, Zacharias II-418
Lemme, Diana IV-698
Leotta, Francesco IV-637
Lepreux, Sophie I-211
Lewis, Richard I-193
Li, Haidong V-82
Li, Yan IV-568
Li, Yueqing IV-594
Li, Yu-Wen III-359, IV-177
Liang, Chao V-171, V-206
Liang, Xueyi V-82
Lietz, Holger V-196
Lietzenmayer, Ryan IV-135
Lim, Brian Y. V-92
Lim, Ji Hyoun IV-368
Lima, Gabriel Vial Correa II-3
Lima, Tania II-428
Limbrecht-Ecklundt, Kerstin V-301
Limongelli, Carla II-434
Lin, Ming-Hui IV-506, V-483
Lin, Qing-Wen III-193
Lin, Ray F. III-199
Lin, Yuan-Pin V-448
Lin, Yu-Ting V-483
Lindley, Craig V-403
Liu, Cheng-Li III-475
Liu, Hai-Lin V-82
Liu, Rong V-206
Liu, Shijing IV-561, IV-618
Liu, Yijun V-498
Liu, Yikun IV-658
Løvborg Jensen, Kasper III-332
Lombardi, Matteo II-434
Long, Zhiying V-498
Longo, Lucia III-277
Lorite, Salvador I-231
Lou, Jian I-555
Lu, Ke V-153, V-216
Lucid, Brian IV-668

Luderschmidt, Johannes I-381
Luo, Bin V-227

Ma, Tao IV-291
Ma, Wenqi IV-551
Macek, Tomáš IV-59
Maekawa, Yasuko II-134
Majewski, Martin IV-147
Majima, Yukie II-134
Makedon, Fillia IV-43
Maki, Atsushi V-411
Malagardi, Ioanna IV-13
Manssour, Isabel H. III-117
Manthey, Robert V-196
Marani, Alessandro II-434
Marcus, Aaron I-13
Marinc, Alexander IV-147
Märtin, Christian I-300, I-340, I-401
Martin, Jennifer L. I-166
Martin, Jim III-97
Martinie, Célia I-51, I-290
Martins, Paulo V-39
Massey, Adam II-36, II-189
Matsui, Tatsunori IV-398
Matsumoto, Shimpei II-444
Matsuura, Kenji II-84
Matsuyama, Takashi IV-408
Mattar, Nikita V-102
Mayas, Cindy I-85, I-391
Mbatha, Blessing II-454
Mbatha, Mbali II-454
McDougall, Siné I-575
Mecella, Massimo IV-637
Medjkoune, Sofiane IV-77
Meixner, Gerrit I-411
Mele, Maria Laura I-203
Mello, Erick II-197
Mendori, Takahiko II-521
Meneses Viveros, Amilcar IV-301
Mennecke, Brian III-428
Mhd Pauzi, Muhammad Asyraf IV-750
Micheals, Ross J. IV-449
Miike, Katsuaki I-565
Miles, Robert II-36
Min, Wensheng IV-540
Minker, Wolfgang III-59
Miranda, Pedro Augusto da Silva e Souza II-3
Mita, Yuusaku IV-416
Mito, Kazuyuki IV-225
Mitsuishi, Takashi I-565
Miwa, Shotaro V-178
Miyaji, Chikara II-84
Miyaji, Yutaka I-620
Mizuno, Tota IV-225
Modesto, Débora Maurmo III-324
Mohagheghi, Amir II-76
Mollard, Régis IV-429
Mondin, Fabio Luciano III-277
Mont'Alvão, Claudia Renata II-127, II-144, II-370
Monteiro, Ingrid IV-439
Morandini, Marcelo I-184, I-249
Moreau, Guillaume IV-530
Moreira, Samuel I-421
Moriwaka, Makoto IV-320
Morreale, Patricia II-335
Morrissey, Kellie IV-87
Motta, Gustavo H.M.B. II-197
Mouchere, Harold IV-77
Mourlas, Constantinos II-418
Mugellini, Elena II-531, IV-157, IV-167, V-19, V-186
Munemori, Jun III-174
Muñoz, Adolfo I-231
Murakami, Satoru I-565
Murata, Atsuo IV-312, IV-320, IV-330
Murata, Kazuyoshi IV-462, IV-491

Nagano, Hiromi V-319
Nakagawa, Koji II-84
Nakamura, Tetsuaki IV-126
Nakamura, Yutaka III-156
Nakatani, Yoshio III-295, V-329
Nakatsu, Ryohei IV-601
Nam, Chang S. IV-561, IV-594, IV-610, IV-618
Neris, Vânia Paula de Almeida II-229, III-227
Nesi, Paolo III-259
Nguyen, Vinh-Tiep IV-67
Nguyen-Huynh, Duy-Hung IV-513
Nicholas, Michael II-640
Nieminen, Marko II-101
Nishida, Shogo III-314, IV-378, V-421
Nishino, Kazunori II-511
Nishiuchi, Yusuke II-521
Noguchi, Anna IV-469
Noor, Nor Laila Md. II-376
Norcio, Anthony F. II-173

Nothdurft, Florian III-59
Nouri, Elnaz II-266
Nouri, Jalal II-464
Novick, David IV-97
Novotny, Philipp V-454
Nunes, Fátima de Lourdes dos Santos I-184

Oami, Takuma IV-568
Obata, Akihiko III-307
Obata, Yuhki IV-568
Oblaender, Vera IV-706
Ochi, Keita IV-330
Oe, Tatsuhito IV-469, IV-712
Ogawa, Hitoshi III-20
Oh, Hyung Jun I-103
Ohiro, Tomoya V-329
Ohkawa, Yuichi I-565
Ohkura, Michiko I-585, II-343, II-647
Ohmori, Kosuke V-421
Ohori, Kotaro III-307
Okada, Akira V-353
Okada, Naoki III-314
Okawa, Jun V-372
Oki, Maho III-209
Okuuchi, Keita IV-340
O'Malley, Claire II-376
Ono, Yuichi II-474
Orfgen, Marius I-411
Otmane, Samir IV-530
Owen, Charles B. V-464
Özcebe, Esra I-310
Ozturk, Elif II-408

Palanque, Philippe I-51, I-290
Pang, Natalie II-181
Paolucci, Michela III-259
Pappu, Aasish IV-107
Paredes, Hugo V-39
Park, Jae Heon III-402
Park, Jongsoon III-217
Park, Kyoung Shin V-339
Park, Kyung Eun II-56
Park, Sangin V-346, V-441
Parthasarathy, Rangarajan II-150
Pastor, Enric I-231
Paternò, Fabio I-241
Paulick, Peyton II-163, V-454
Peeters, Jeroen I-128
Peng, Xiaobo II-408
Pengnate, Supavich I-593
Perez Pelaez, Mariano IV-348
Peters, Anicia III-428
Petersson Brooks, Eva III-418
Petitrenaud, Simon IV-77
Pfaff, Mark S. IV-658
Pham, Truong-An IV-67
Philippow, Ilka I-510
Pichelmann, Stefan II-612
Pierantoni, Felipe II-127, II-144
Pino, Alexandros IV-358
Platt, Donald IV-629
Plischke, Herbert II-163, V-454
Ponsa, Pere I-231
Porras, Jari II-239, V-431
Potamianos, Gerasimos IV-43
Prates, Raquel Oliveira I-3
Price, Chandler III-285

Rabe, Felix IV-117
Rabie, Osama II-173
Rahmanivahid, Pooyan II-76, V-262
Raine-Fenning, Nicholas II-36, II-189
Rampoldi-Hnilo, Lynn II-309
Ramsbrock, Jens II-621
Randall, Tania I-109
Raposo, Alberto IV-439
Rauch, Marta II-276
Rauff, Stefanie III-438
Ray, Jerry III-285
Rebenitsch, Lisa V-464
Rehm, Matthias I-431, II-361, III-332
Rekimoto, Jun V-112
Reppa, Irene I-575
Riccio, Angela IV-637
Ridi, Antonio IV-157
Rigas, Dimitrios IV-23
Rintel, Sean III-484
Ritter, Marc V-196
Rivera, Ismael V-122
Robichez de Carvalho, Gustavo II-3
Rocha, Mario Sandro II-428
Rodeiro-Iglesias, Javier I-441, I-490
Rodil, Kasper III-332
Rodrigues, Kamila III-227
Røssvoll, Till Halbach III-68
Romano, Marco III-342
Roselli, Teresa II-484
Rossano, Veronica II-484
Rudnicky, Alex IV-107

Ruffieux, Simon V-186
Rukavina, Stefanie V-474
Ryu, Taebeum IV-368

Saito, Yukou II-647
Sakairi, Takeo IV-722
Sakamoto, Kiyomi V-353
Sakamoto, Maki IV-126
Sakamoto, Takehiko IV-730
Sakamoto, Yuichiro IV-469
Sakashita, Seiji V-353
Sakata, Nobuchika IV-378
Sakoda, Masayuki II-134
Sakurai, Yoshihisa II-84
Samaras, George II-418
Sampanes, Anthony Chad II-284
Sampanthar, Kes II-292
Sánchez, Jaime II-299
Sandnes, Frode Eika II-20
Santos, Oscar Francisco dos I-249
Santos, Thebano Almeida II-428
Sarathy, Rathindra I-593
Sasaki, Shiori II-493
Sato, Ayaka V-112
Sato, Kosuke V-270
Satoh, Hironobu II-521
Scerri, Simon III-29, V-122
Scheruhn, Hans-Jürgen III-446
Schettini, Francesca IV-637
Schieck, Ava Fatah gen. V-10
Schierz, Christoph II-163
Schlegel, Thomas IV-698, V-46
Schlick, Christopher M. III-127
Schmidt, Michael IV-479
Schmidt, Werner I-330, III-456
Schnädelbach, Holger II-36
Schneidermeier, Tim I-176
Scholz, Sebastian C. I-451
Schuller, Andreas I-256, III-29, V-122
Schwaller, Matthias IV-388
Schwarz, Markus J. V-454
Sciarrone, Filippo II-434
Sebillo, Monica III-342
Segers, Mariël III-371
Segura, Vinícius I-460
Seki, Hironori II-647
Seki, Makito V-178
Selvarajah, Kirusnapillai II-36, II-189
Seo, Ki-Young V-339
Servières, Myriam IV-530
Shahid, Suleman II-117, II-541
Shaw, Heather IV-668
Shaw, Terry III-97
Shen, Yang Ting V-132
Sherry, Lance I-193
Shiba, Haruya II-521
Shibuya, Yu IV-462, IV-491
Shiizuka, Hisao I-601
Shikanai, Nao I-611
Shimizu, Yuichiro IV-126
Shin, Yunhee V-362
Shizuki, Buntarou IV-272, IV-416, IV-469, IV-712, V-72
Shoji, Hiroko V-372
Siegert, Ingo V-301, V-381
Siew, Siang-Ting I-470
Siio, Itiro III-209
Silveira, Milene S. I-3, III-117
Simione, Luca IV-637
Simões, Fabiana I-460
Singh, Satinder I-193
Siriwardana, Sanath IV-601
Sloan, David I-41
Snyder, Michele II-309
Soares, Evandro II-428
Sodnik, Jaka II-551
Soga, Masato II-134
Son, Jae Pyo II-631
Song, Bifeng IV-540
Song, Sutao V-491
Song, Yeong-Tae II-56
Sotero, Gabriel I-460
Soui, Makram V-141
Sousa, João IV-678
Souza Filho, Guido II-197
Sowmya, Arcot II-631
Spiegel, Götz III-396
Sridharan, Seshadri IV-107
Srivastava, Sumit IV-497
Stary, Christian I-330, III-456
Steinhoff, Camie I-265
Sterbini, Andrea II-351
Still, Jeremiah D. I-265
Stillwater, Tai II-640
Stocklöw, Carsten IV-147
Storz, Michael V-196
Sturm, Christian III-352
Sugano, Shohei I-620
Sugiyama, Masayuki IV-456
Sun, Ming IV-107

Sung, Tae-Eung V-362
Suzuki, Ayaka IV-469
Suzuki, Ryo IV-348

Tabak, Feride S. II-391
Tajima, Terumasa III-466
Takahashi, Katsumi I-77, I-119, I-137, I-500, I-531
Takahashi, Nobumichi II-647
Takahashi, Shin IV-730
Takano, Kosuke II-493
Takao, Shizuka III-174
Takemura, Noriko V-270
Takeno, Hidetoshi I-159
Takimoto, Munehiro IV-235
Tamada, Takashi IV-722
Tamborello, Franklin P. IV-449
Tamura, Kazuki IV-740
Tan, Jun-Wen V-474
Tanaka, Jiro IV-196, IV-272, IV-281, IV-416, IV-469, IV-712, IV-730, V-72
Tang, Honglin IV-243
Tanikawa, Kyohei III-314
Tano, Shun'ichi I-480
Tauchi, Hikaru II-398
Tavares, Tatiana A. II-197
Tawatsuji, Yoshimasa IV-398
Teixeira, Cesar Augusto Camillo III-227
Teixeira-Faria, Pedro M. I-441, I-490
Temperini, Marco II-351
Theng, Yin-Leng II-181
Thiel, Simon III-29
Thome, Rainer III-411
Tian, Yun V-280
Ting, Chih-Hung IV-506
Togawa, Satoshi II-503
Tokdemir, Gul I-310
Tokumaru, Masataka V-319
Tomimatsu, Kiyoshi IV-568, IV-578
Tomiyama, Ken I-620
Tong, Xin II-207
Torenvliet, Gerard I-109
Tornero, Josep I-231
Tortora, Genoveffa III-342
Tran, Minh-Triet IV-67, IV-513
Traue, Harald C. V-474
Traum, David II-266
Tripathi, Ramesh Chandra IV-497
True, Nicholas I-128
Truong, Chau Thai IV-513
Tsai, Shih-Lung III-78
Tsai, Teng-Yao IV-506
Tsai, Tsai-Hsuan III-359
Tsai, Yi-Chien III-199
Tseng, Steven I-540
Tsianos, Nikos II-418
Tsukada, Koji III-209
Tsumori, Shin'ichi II-511
Tsunoda, Koichi II-398
Tsuruda, Yu V-391
Tung, Fang-Wu III-237
Tung, Tony IV-408
Tzemis, Evangelos IV-358

Uang, Shiaw-Tsyr III-475
Ueda, Kazutaka II-398
Ueda, Yoshihiro I-77, I-119, I-137, I-500, I-531
Uetsugi, Raku IV-312
Ushio, Shuhei II-647

Valla, Massimo V-122
Valverde, Lauralee II-408
Vedhara, Kavita II-36, II-189
Viard-Gaudin, Christian IV-77
Vilimek, Roman II-578, II-621
Viller, Stephen III-484
Vitiello, Giuliana III-342

Wachsmuth, Ipke IV-117, V-102
Wahab, Norshahriah IV-523
Wallach, Dieter I-451
Walter, Steffen V-301, V-474
Wang, Hongmei IV-135
Wang, Jan-Li III-193
Wang, Liang V-206
Wang, Lingyun (Max) I-350
Wang, Pengwei V-254
Wang, Qian V-216
Wang, Weizhong V-280
Wang, Wenzhong V-227
Wang, Yao-lien V-483
Wang, Yijun V-448
Wang, Yixue V-153
Wang, Yuping V-246
Wang, Zhaoqi V-227
Wästlund, Erik III-10
Watanabe, Keita V-112
Watanabe, Shintaro V-178
Webb, Erika Noll II-316

Weber, Gerhard IV-479
Weber, Julian II-621
Wee, William IV-291
Wei, Chun-Shu V-448
Wei, Ming-Hsuan II-654
Weigang, Li IV-648
Wen, Chao-Hua III-78
Wendemuth, Andreas V-301, V-381
Wendler, Stefan I-510
Wenzel, Christopher II-46
Whang, Mincheol V-346, V-441
Wilcock, Graham IV-32
Winckler, Marco I-51, I-521
Winschiers-Theophilus, Heike III-332
Wolff, Christian I-176
Wolff, Marion IV-429
Wong, Chui Yin IV-750
Wong, Man Leong IV-750
Woodjack, Justin II-640
Worpenberg, Annika II-323
Wu, Guanshang II-207
Wu, Qiong II-207
Wu, Siju IV-530
Wu, Tong-Ying II-20
Wu, Zhanwei I-555
Wu, Zhongke V-280

Xiaohui, Tan V-289
Xiong, Chenlin V-254
Xu, Bingxin V-236
Xu, Lele V-498
Xuesong, Wang V-289

Yabe, Takao II-398
Yajima, Hiroshi II-216
Yalvac, Bugrahan II-408
Yamaguchi, Takehiko IV-561
Yamaguchi, Takumi II-521
Yamane, Shohei III-307
Yamaoka, Toshiki I-361
Yamashiro, Mitsuo II-474
Yamashita, Kuniko V-353
Yamazaki, Kazuhiko I-77, I-119, I-137, I-500, I-531
Yan, Fu V-289
Yanagida, Koji I-77, I-119, I-137, I-500, I-531
Yang, Chao-Yang II-20
Yang, Jason III-484
Yang, Lei V-216
Yang, Weixin V-254
Yang, Yifang V-246
Yang-Mao, Shys-Fan V-483
Yao, Li V-491, V-498
Yasu, Hiroko I-146
Yasumura, Michiaki V-112
Yazdifar, Mahshid II-76, V-262
Yazdifar, Mohammadreza V-262
Yeo, Alvin W. I-470
Yildiz, Alparslan V-270
Yildiz, Ekrem I-310
Yin, Yabo V-280
Yoon, Joonsung I-32
Yoshida, Masanobu II-521
Yoshikawa, Takuto IV-416, IV-469
Yoshikawa, Tomohiro IV-740
Yoshimura, Hiroki V-391
Yu, Der-Jang I-540
Yu, Jiamin I-555
Yue, Yong V-19
Yun, Myung Hwan IV-368

Zaki, Michael I-340
Zaman, Tariq I-470
Zeng, Wen-Jun IV-506, V-483
Zetali, Karwan II-464
Zhan, Yu V-491
Zhang, Hang V-498
Zhang, Jiacai V-491
Zhang, Rushao V-498
Zhang, Xin V-254
Zhang, Yugang IV-540
Zhao, Fan II-46, III-492
Zhaolao, Lu V-289
Zhong, Zhengyang V-254
Zhou, Mingquan V-280, V-289
Zhou, Xuefu IV-291
Ziefle, Martina III-49