Chi sparks 2014

Jos van Leeuwen, Pieter Jan Stappers,
Maarten Lamers, Maarten Thissen (Eds.)

Creating the Difference

Proceedings of the Chi Sparks 2014 Conference

HCI Research, Innovation, and Implementation

The Hague University of Applied Sciences
The Netherlands
April 3, 2014

chi Nederland

THE HAGUE UNIVERSITY OF APPLIED SCIENCES

Chi sparks 2014

Creating the Difference

Proceedings of the Chi Sparks 2014 Conference

The Hague University of Applied Sciences
The Netherlands

April 3, 2014

HCI Research, Innovation, and Implementation

Chi Sparks 2014

Editors

Jos P. van Leeuwen

The Hague University of Applied Sciences
j.p.vleeuwen@hhs.nl

Pieter Jan Stappers

Delft University of Technology
p.j.stappers@tudelft.nl

Maarten H. Lamers

Leiden University
m.h.lamers@liacs.leidenuniv.nl

Maarten J.M.R. Thissen

The Hague University of Applied Sciences
m.j.m.r.thissen@hhs.nl

Published by:

The Hague University of Applied Sciences, The Netherlands

and

Chi Nederland, the Dutch association promoting Human-Computer Interaction in The Netherlands (www.chi-nederland.nl)

ISBN: 978-90-73077-55-3

These proceedings are available in the public domain at:
http://proceedings.chi-sparks.nl

ORGANISATION

The Chi Sparks Conference 2014 was organized by Chi Nederland and The Hague University of Applied Sciences. It was hosted at The Hague University of Applied Sciences, Academy of ICT & Media, on April 3, 2014.

Conference Chairs

Jos van Leeuwen, Pieter Jan Stappers, Maarten Lamers, Maarten Thissen

Management Team

Jannie Oosting, Leon Schröder, Stefan Mirck

Organising Committee

Alex Reuneker, Bram Reurings, Cecilia Tan, Danica Mast, David Dielen, Ellen Grummels, Elvira Arslanova, Fons Verbeek, Gaby Eenschoten, Iefke Bloothoofd, Koen van Turnhout, Liou Yamane, Lisanne Groenendaal, Lynda Hardman, Nick van 't End, Patrick Deters, Petra Visser, Roald Dijkstra, Sanne Verbaan, Selma Boenders, Stephanie van der Meer, Vanessa Vakili

Scientific Committee

Aadjan van der Helm, Alex Reuneker, Alwin de Rooij, Amalia Kallergi, Arnold-Jan Quanjer, Ben Kirman, Bram Reurings, Carla Hoekendijk, Caroline Maessen, Conor Linehan, Danica Mast, David Geerts, Dennis Dams, Dick Rijken, Dries de Roeck, Edwin van der Heide, Erik de Vries, Fons Verbeek, Geert de Haan, Ianus Keller, Jan-Dirk Schagen, Jane Sattoe, Jelle van Dijk, Joey van der Bie, Joost Burger, Jos van Leeuwen, Koen van Turnhout, Lenny van Onselen, Lilli Sukirman, Lynda Hardman, Maarten Lamers, Maarten Thissen, Marc Steen, Marie de Vos, Martijn Hartog, Mascha van der Voort, Mathieu Gielen, Matthias Rauterberg, Myriam Traub, Patrick Kruithof, Peter Becker, Pieter Jan Stappers, Remko van der Lugt, Robbert Winkel, Ru Klein, Shaun Lawson, Suzanne Hallenga-Brink, Theo Rooden, Thijs Waardenburg, Valentijn Visch, Vincent Vijn

Table of Contents

Creating the Difference - Editorial **7**
Jos P. van Leeuwen, Pieter Jan Stappers, Maarten H. Lamers, Maarten J.M.R. Thissen

Technology Push **10**

Illuminating Shadows: Introducing Shadow Interaction in Spatial Augmented Reality 11
Marcello Gómez-Maureira, Carolien Teunisse, Hanna Schraffenberger, Fons Verbeek

Auditory Signal Design for Automatic Number Plate Recognition System 19
Chris Heydra, Reinier Jansen, René van Egmond

Presence Mediator: Creating Social Presence through Spatial Positioning of Everyday Objects 24
Daniel Egerer, Jelle van Dijk

Automated Mobile User Experience Measurement: Combining Movement Tracking with App Usage Logging 31
Lucas Noldus, Ben Loke, Mike Kelia, Andrew Spink

Where We Live **35**

Family in Focus: On Design and Field Trial of the Dynamic Collage [DC] 36
Jasper Jeurens, Koen van Turnhout, René Bakker

Creating Better Health Worker - Patient Interaction Using ICT; Design for Applicability and Acceptance 45
Arthur Bennis, Dick Lenior

Experience Centered Design of Energy Interventions for Shared Student Accommodation 50
Conor Linehan, Derek Foster, Shaun Lawson, Maureen Schoonheyt, Katrin Heintze

Measuring Social Media Activities of Dutch Museums: Developing a Social Media Monitor 60
Thijs Waardenburg, Rogier Brussee, Erik Hekman

User Perspectives **64**

Dynamic Layering Graphical Elements For Graphical Password Schemes 65
Wouter van Eekelen, John van den Elst, Vassilis-Javed Khan

Persuasive by Design: a model and toolkit for designing evidence-based interventions 74
Sander Hermsen, Reint Jan Renes, Jeana Frost

Pointing Task Evaluation of Leap Motion Controller in 3D Virtual Environment 78
Joanna Camargo Coelho, Fons Verbeek

How Turn-Taking Influences the Perception of a Suspect in Police Interviews 86
Rifca Peters, Merijn Bruijnes, Rieks Op Den Akker

Kids @ Play **90**

Wobble: Supporting Social Play through an Open-ended Play Environment 91
Alice van Beukering, Linda de Valk, Tilde Bekker

BioPong: Adaptive Gaming Using Biofeedback 100
Dirrik Emmen, Georgios Lampropoulos

Digital Learning Environments and Collaborative Pedagogy: Media Culture 2020 104
Martyn Thayne, Graham Cooper

The Kids' Knowledge Base: Connecting Junior Science to Society 108
Aldo de Moor

Getting It Done 112

'Include', a Toolbox of User Research for Inclusive Design 113
Hanna Zoon, Anita Cremers, Berry Eggen

Breaking Usability Rules to Enable Reflection 122
Arnold Quanjer, Maarten Lamers

Ambient and Pervasive Design: HCI Education from UbiComp to Creative Design 126
Geert de Haan

Building UX capabilities at a technology firm: A shared success story 130
Jesse Grimes, Anneke van de Langkruis

Concept Plan for User Testing in Scrum Projects 133
Sirvan Ugur, Swie Oei, Arjen de Bruijn

Demos & Posters 138

Giving user-generated content back to the users: Testing data de-centralisation with active content creators 139
Amy Guy

Make Your Own HCI! 142
Annika Geurtsen, Bobbie Smulders

Digicoach: a digital coach for self-tracking athletes 145
Antoine Hogenboom

Iris Intranet: Enhancing the Way People Work 147
Arthur Turksma, Jos Rouw

ixi-Play, a robot buddy for young children 149
Bart Dirkx, Ruud van de Aalst

Interactive Visualization of Video Data for Fish Population Monitoring 152
Emma Beauxis-Aussalet, Lynda Hardman

"Graham (aged 8)" Memory box demo 156
Graham Cooper

The AffectButton: a Digital Self-report Tool for Emotion 158
Joost Broekens

ReSearchNow – Playful App for Participation and Engagement in (Design) Education 160
Marije ten Brink

Gamesourcing Expert Painting Annotations 163
Myriam C. Traub, Jacco van Ossenbruggen, Jiyin He, Lynda Hardman

The Programming Language as Human Interface 166
Steven Pemberton

REHAP – Modular and Interactive Rehabilitation Tiles 169
Victor Donker, Bert Bongers, Panos Markopoulos

Author Index 172

Creating the Difference
Editorial

'Creating the Difference' is the theme of the 2014 edition of the Chi Sparks conference. It is also the challenge that the Human-Computer Interaction (HCI) community is facing today. HCI is a creative field where practitioners engage in design, production, and evaluation of interactions between people and digital technology. Creating excellent interfaces for people, they make a difference in media and systems that people are eager to use. Usability and user experience are fundamental for achieving this, as are abilities at the forefront of technology, but key to a successful difference is getting the right concepts, addressing genuine, intrinsic, human needs. Researchers and practitioners contribute to this area from theory as well as practice by sharing, discussing, and demonstrating new ideas and developments. This is how HCI creates a difference for society, for individuals, businesses, education, and organizations.

The difference that an interactive product or service makes might lie in the concept of it but also in the making, the creation of details and the realisation. It is through powerful concepts and exceptional quality of realisation that innovation is truly achieved.

At the Chi Sparks 2014 conference, researchers and practitioners in the HCI community convene to share and discuss their efforts on researching and developing methods, techniques, products, and services that enable people to have better interactions with systems and other people. The conference is hosted at The Hague University of Applied Sciences, and proudly built upon the previous conferences in Arnhem (2011) and Leiden (2009).

KEYNOTE LECTURES

The conference enjoys the presence and performance of five keynote lecturers. Thomas Marzano, Global Head of Brand Communication Design at Philips Design, challenges the HCI community to think about Brand Experience instead of User Experience. "Brand Experience, there's no app for that…" Tapping from his experience with the new Philips Brand he shows us how a company should approach its brand in a holistic way and thus creating a better and deeper felt brand differentiation.

Ohyoon Kwon and Albert Kivits give us "The making of HomelessSMS," a story of how designers, technologists and entrepreneurs work with homeless people, service professionals and directors of organisations. They collaborated to provide social and economic added value enabled by everyday mobile technology in a complex social problem - homelessness. This talk also reflects how this social innovation project has been evolved ever since 2010 in East London initiated by Will Brayne and has gone through several iterations across South Korea and The Netherlands, spinning off a new business activity.

STEIM has been at the forefront of new interfaces for musical expression since 1969, and the presentation will feature some of the work that has been done in the past. Dick Rijken and Frank Baldé demonstrate how they have come a long way since then, with "Music, Intuition, and Interfaces." New technologies have not only changed the things we can make, but also how we make them, and, increasingly: why we make them. Today's complex problems are forcing us to rethink the relation between thinking and intuition, between the mind and the body, and between control and surprise. Musical instruments are a good context for experimenting with these issues. Dick and Frank outline current developments in the world of music and musical instruments and interfaces, and also offer a live performance using new technology from STEIM that is available also to designers of interactive instruments and installations.

CONTRIBUTIONS FROM THE HCI COMMUNITY

In addition to the three excellent keynote addresses at the conference, these proceedings offer 21 contributions, from academia as well as industry, which we have organised in five sections.

TECHNOLOGY PUSH

While new technologies for interaction emerge from the drawing boards and laboraties of engineers, it is the joyful task of interaction designers to create useful applications of new technology that have meaning and relevance to people. It is by forging technical capabilities into useful and usable products and services that innovation offers its true value for society.

Gómez-Maureira, Teunisse, Schraffenberger, & Verbeek use shadows both as an interaction input and as an area for display in spatial augmented reality. This allows them to create an environment for physical interaction with information, useful for, e.g., augmenting museum exhibits with information. (p.11)

Heydra, Jansen, & van Egmond designed an auditory signal for a system used in police cars to automatically recognise suspicious number plates. The signal design takes the environment in the police car into account and helps identify and locate the suspicious car in the vicinity of the police car. (p.19)

Social interactions between people living together are reflected in the rearrangement of objects they use together. **Egerer & van Dijk** experimented with simulating this phenomenon in long-distance relationships. They discuss the possibilities and effects of synchronizing the position of domestic objects in remote places in order to mediate social presence. (p.24)

Measuring the interaction between user and system is an essential part of good interaction design. **Noldus, Loke, Kelia, & Spink** present a set of tools for user experience studies that automates this on mobile devices, visualising the user's activities on the device as well as their location and spatial behaviour, over extended periods of time. (p.31)

WHERE WE LIVE

Innovation has the strongest impact when it takes place closest to us, at our homes and in our daily lives. In this section we collected four papers that offer experiences and insights from projects that innovate the places we live and the ways we live together with others.

Informal care of elderly people, by their relatives, is increasingly important in our ageing population. **Jeurens, van Turnhout, & Bakker** describe their design process of a system that encourages and facilitates social involvement of family members in the care of elderly people. It motivates social awareness by connecting physical presence of one relative with attention and involvement on a distance of others. (p.36)

Acceptance is one of the key aspects of user experience. **Bennis & Lenior** discuss three case studies of so-called telecare applications, in order to investigate what user-centred design approaches may lead to better acceptance of such tools for varying tasks and circumstances. (p.45)

Linehan, Foster, Lawson, Schoonheyt, & Heintze present an experience-centred design of interventions to motivate reduced energy consumption in pre-paid rented accommodations. The paper describes their research and design method, based on large-scale participatory design, that aimed to elicit experiential and reflective data in order to inform the design process by identifying experiences, perceptions, attitudes, behaviours, challenges, and opportunities. (p.50)

Social media play an increasingly important role in the online strategy of organisation. In order to assess the communication through social media of such organisations as museums, **Waardenburg, Brussee, & Hekman** have developed a monitoring system that is meant to continually track social media activities of the heritage sector and mine its history as well. (p.60)

USER PERSPECTIVES

User-centred design is an old adage in the HCI field that places the user's needs, abilities, and objectives at the centre of the design process. From a user's perspective, technology will work only when it contributes to the needs and objectives, and leverages the abilities. This section introduces four projects where the focus on user perspectives drives the design of new tools.

Van Eekelen, van den Elst, & Khan propose a framework for classification of graphical password schemes. In doing so, they also discover that a combination of graphical elements, e.g., shape and colour, can be used to devise password schemes that are both easier to memorise and less prone to shoulder attacks. (p.65)

More and more, the objective of a designed interaction is to change the behaviour of its users. Knowledge of psychological theory is essential in achieving desired results, but often not sufficiently available to designers. **Hermsen, Renes, & Frost** present a tool that help designers in creating evidence-based interventions for behavioural change. (p.74)

Navigating, pointing, and manipulating objects in a virtual environment has been subject of research for many years now. With new interaction devices coming to the market, optimal ways of interacting in virtual environments are still to be determined. **Coelho & Verbeek** experimented with the Leap Motion sensor that is capable of sensing hand gestures and compare its performance to that of other 3D input devices. (p.78)

Serious games are an effective tool in training skills and behaviour. **Peters, Bruijnes, & Op Den Akker** contribute to this area, focusing on social skill training and argue that the credibility of virtual agents in serious games is crucial for this purpose. Their paper reports on how appropriate turn-taking in conversations by such agents influences their credibility in the game. (p.86)

KIDS @ PLAY

Games are important to kids, not just for the fun of it, but also because they learn from them. They acquire mental and motor skills as well as social skills by playing games, learn solitary or in groups, whether collocated or at a distance.

Social play is essential for the development of children and the design of a play environment that encourages social behaviour is important. **Van Beukering, de Valk, & Bekker** developed an open-ended play environment that allows children to engage in various levels of social activity and that supports the three stages of social play: invitation, exploration, and immersion. They discuss how various elements of interaction design have helped develop this play environment. (p.91)

With the abundant availability of sensors in our lives, including those on our bodies, gameplay is being taken to the next level. **Emmen & Lampropoulos** study the possibilities to use physiological data, such as heartbeat and galvanic skin response, to adapt game attributes for the individual user's present state. They argue that this kind of data can be used to assist gamers during gameplay. (p.100)

Thayne & Cooper present a set of digital tools that were developed for collaborative teaching in the Media Culture 2020 programme. This involves a range of social media platforms and open, virtual environments for collaborative learning. When properly and effectively utilised, these tools are not merely a vehicle for distant collaboration in learning, they also help establish trust and bonds between students and teaching staff. (p.104)

De Moor presents an instrument that helps rouse the interest of young children in science and technology. The Kid's Knowledge Base forms the basis for collaborative activities in a network of educational stakeholders, from primary schools to universities, utilising a combination of online and physical tools. Although the tools are important, de Moor concludes that it is essential in each context to find the right connections between content, tools, activities, and stakeholders. (p.108)

GETTING IT DONE

HCI is an intriguing domain with unique challenges, because of the complexity of the technology, the variety of the application domains, the wide and often conflicting interests of the stakeholders involved, the incredible speed of development, and not in the least because of the human impact. This forces us to reflect and critique our methods and to incessantly search for better tools, techniques, approaches, guides, principles, and criteria for designing, producing, and evaluating our designs.

Inclusive design is the area of HCI where we aim to empower users through products and services that match their capabilities. **Zoon, Cremers, & Eggen** offer a toolbox, consisting of an app and a book, that helps designers in practice to find suitable research methods for a range of specific groups of users, such as elderly people, or people with low literacy. (p.113)

Human-computer interactions are often optimised towards effectiveness and usability. There are situations, however, when an interface should encourage reflection before interaction occurs. **Quanjer & Lamers** argue that the rules of usability may need to be broken to instil reflection in users, for example by enforcing a delay in the interaction to make sure that actions are purposeful. (p.122)

De Haan discusses the design of a course on ambient and pervasive design, in a curriculum on human-centred creative technology. Topics such as the Internet of Things and Ubicomp form parts of this course. The challenge for educators is to embed these topics and research skills in these areas without becoming too academic. (p.126)

Grimes & van de Langkruis report on a successful effort to establish the discipline of user experience in a company that was traditionally technology-focused. Their approach was to introduce the UX awareness from strategy down to operations, gradually building the capabilities in the company that were previously fully outsourced. (p.130)

Scrum brings an interdisciplinary team of designers and developers together with the product owner and project manager, in an agile process of designing and developing a new product. Testing with users is a particular challenge in this rapidly moving context. **Ugur, Oei, & de Bruijn** propose a set of guidelines to integrate user testing in scrum projects. This approach will help address usability issues during every sprint in the project, allowing for improvements in the next sprint based on test results. (p.133)

IN CONCLUSION

The 2014 edition of the Chi Sparks conference has brought together a fascinating blend of research output and practical case studies from both industry and education. It proves once again how wonderfully broad the HCI domain is, with many areas of application, and in strong relation with life itself.

The contributions to the conference, gathered in these proceedings, inspire and motivate researchers and practitioners alike to continue the quest for improving our methods, our results, and indeed our goals.

As programme chairs of the conference, we were pleased to notice an increase in the participation of students as authors in the conference. This tells us that educators actively stimulate students to participate in research activities and to publish on their results. We regard this as an essential development for expanding the domain's power to innovate, both in research and practice.

The contributors and audience of the conference have beautifully adopted the conference theme, 'creating the difference,' by showing us their efforts to find and develop methods and means to create the right concepts, and to apply innovative technologies in a sensible and sensitive way.

The conference offered a stage not only to the 21 papers presented in these proceedings, but also to demos and posters displayed at the conference. The short papers accompanying the demos and posters are included at the end of these proceedings.

Finally, we would like to express our gratitude to all the volunteers who made this conference possible, including the many reviewers that participated in selecting and improving the contributions.

The Hague, April 3, 2014
Jos van Leeuwen, Pieter Jan Stappers,
Maarten Lamers, Maarten Thissen
Conference Chairs

TECHNOLOGY PUSH

Illuminating Shadows: Introducing Shadow Interaction in Spatial Augmented Reality

Marcello A. Gómez-Maureira
LIACS, Leiden University
Niels Bohrweg 1, Leiden, NL
ma.gomezmaureira@gmail.com

Carolien Teunisse
LIACS, Leiden University
Niels Bohrweg 1, Leiden, NL
lucidlien@gmail.com

Hanna Schraffenberger
LIACS, Leiden University
Niels Bohrweg 1, Leiden, NL
h.k.schraffenberger@liacs.leidenuniv.nl

Fons Verbeek
LIACS, Leiden University
Niels Bohrweg 1, Leiden, NL
f.j.verbeek@liacs.leidenuniv.nl

ABSTRACT

In this paper we present a new mode of interaction in 'Spatial Augmented Reality' (SAR) setups, using shadows as interaction input as well as display area. We claim that the combination of shadow interaction and SAR offers a novel, enjoyable and interesting way of interacting with information in a physical manner. This is especially relevant for contexts such as museum exhibits, where digital information and physical objects relate to one another. The results of our usability experiment with a zebrafish model show that users enjoy the combination of shadow interaction and SAR, as well as see a use for it in exhibition environments.

Author Keywords

Spatial AR, Shadow interface, Public display, Shadow interaction, Spatial interaction, Multi-user awareness, Exhibition display, Emotional display, Playful interaction.

ACM Classification Keywords

H.5.1 Information Interfaces and Presentation: Multimedia Information Systems – Artificial, augmented, and virtual realities;

H.5.2 Information Interfaces and Presentation: User Interfaces – Input devices and strategies.

INTRODUCTION

In the last few years, the use of spatial augmented reality (SAR) [2] – often called 'projection mapping' – has emerged as a popular display method for a variety of uses. Examples range from uses in modern art installations, to live visual performances (commonly referred to as 'VJ-ing'), and augmented prototyping [17]. In spatial AR displays, video projectors are used to display virtual content on physical objects in such a way that the content can be interpreted as part of the physical world. It can be argued that this type of display is most effective when the line between the physical and virtual is blurred. As such, any visual artifact in the projection that separates the virtual from the physical would be detrimental to visual illusion that is created. An example of such artifacts is the shadows that occur if the projection path is occluded.

In our study we explore the idea that instead of being problematic, shadows could be used as interaction method and as a way to define an independent secondary display area in spatial AR setups. We propose a system that allows users to interact with an exhibit by casting shadows on the physical model and thereby changing the displayed content in the shadow areas. We call such a mode of operation 'Shadow Spatial AR Interaction' (SSARI). The occurrence of shadows is a concept that is strongly linked to the physical environment. By using shadows as integral part of an interface we argue that the interaction method can contribute to the spatial AR metaphor; the convergence of the virtual and the physical world. We further argue that the use of SSARI promotes spatial exploration of displayed content in a playful way, thereby involving gameful interaction design methodologies (often referred to as 'gamification') [3].

Given that SSARI combines a physical model with virtual content, we expect that it is particularly suited for museums and exhibition environments that use both physical models and digital information in their exhibits. Our primary research question is therefore: *Is the use of 'Shadow Spatial AR Interaction' (SSARI) a suitable method to present interactive content in museums?*

To elaborate on this concept, and to study the feasibility of SSARI setups in general, we developed a prototype of an over-life sized zebrafish model as projection canvas (cf. fig. 1). The prototype was then evaluated in a user assessment experiment and compared to an alternative interactive setup: a touch-based tablet display (Apple iPad, 2010). This comparison was deemed relevant to investigate as previous studies [7, 14, 18] have described the potential benefits of touch-based interaction systems for presenting museum content.

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

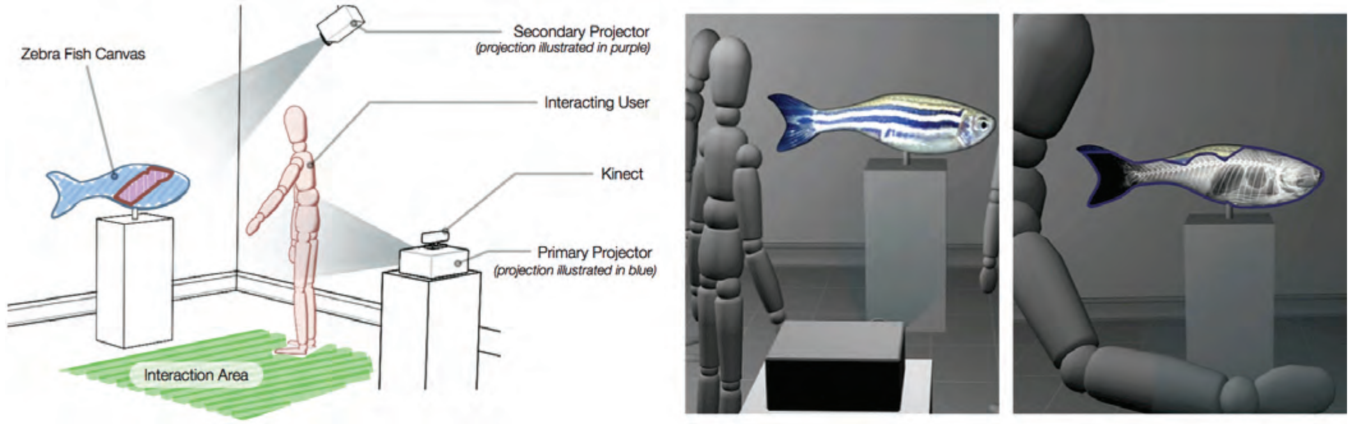

Figure 1: (Left) Schematic illustration of the SSARI setup, (Middle) Primary projection displaying zebrafish skin, (Right) Secondary projection showing X-ray visualization inside shadow area.

Since the nature of interaction is different between the two setups, we compare general use parameters such as ease of use, user enjoyment, suitability in museums, and how informatively the content is presented. Our secondary research question is then: *How does the use of SSARI (in the context of museum displays) compare to the use of touch-based displays?*

With respect to our research questions, we hypothesize that: Users will find the SSARI setup suitable for use in museums (H1). Users will find the SSARI setup easier to use than the tablet (H2). Users will report more enjoyment when using the SSARI setup than when using the tablet (H3). Users will find the SSARI setup more informative in regards to the presentation of content than the tablet (H4). Users will find the SSARI setup more suitable for use in museums than the tablet (H5).

In the following sections of the paper we will discuss important prior works that either involve shadows as interface elements or spatial AR setups. We will then describe the conceptual design of a SSARI setup and the functional design of our zebrafish display that uses it. Subsequently, we describe the setup, execution, and results of our usability assessment efforts before drawing conclusions. Finally, we will discuss the outcome of our study before ending with considerations regarding future work.

BACKGROUND AND RELATED WORK

The combination of SAR with shadow interaction has so far remained unexplored. Previous work has, however, investigated these concepts individually and shown various ways in which they can be used as effective components in user interfaces.

SAR displays are often used to provide unique visual aesthetics that could not be achieved with traditional screens. One research [12] uses two projectors to augment a physical model of the Taj Mahal. By modifying the projection content, the model's physical appearance can be modified in an instant despite the static nature of the model itself. Another study used SAR for medical training purposes: In 'BodyExplorerAR' [13] a full-body mannequin serves as canvas for projection content. Users of the interface can explore the anatomy and physiology of a human, and receive direct feedback to see the consequences of actions they perform on the mannequin.

What these studies have in common is the use of SAR for the purpose of real-time interactions with physical objects. These studies illustrate how useful SAR can be in educational contexts by allowing the easy and dynamic linking of digital information to physical structures. In each of these studies, the occurrence of shadows would introduce an unwanted artifact into the interaction and possibly distract the user's immersion. However, studies that explore shadows as interaction method show that shadows can also be used effectively as part of the interface.

Such an implementation can be seen in a study about crowd audience participation [8] in which a modified version of the game 'Missile Command' is projected onto an elevated screen. The audience can interact with the game by occluding the projection path by hitting a beach ball into the air. Another study introduces the term 'Shadow Reaching', an interaction technology using the properties of perspective in shadow projection to let a single user extend his or her reach on a large screen [15].

Within the context of user interaction, shadows are widely regarded as familiar and intuitive to users, and are therefore considered to be well suited to act as interface element in interactive visual applications [4, 8, 15]. The general understanding of how shadows work offers perceptual advantages for applications that link physical actions to virtual responses and vice versa [9, 20, 15]. Studies [1, 20] have also emphasized favorable economic aspects of using shadows as form of interaction. In addition to such practical considerations, user experiments [1, 9, 6] showed that the involvement of shadows can introduce emotional aspects to an interface due to its expressiveness, inducing a sense of human presence in a virtual environment.

The interaction prototype described in this paper aims to capture and combine the separate benefits of SAR and shadow interaction, specifically visual aesthetics, immediacy, and emotional impact. We believe that the resulting combination would be well suited for interactive exhibitions. Research into interactive museum installations suggests that content should be presented in a way that lets users engage and experience it [5]. This approach is often chosen when creating exhibitions for children, but rarely implemented for older target groups. In our study we aim to develop an interaction prototype that offers adult audiences a way to engage with multiple information layers without compromising the playful experience. We argue that our approach will not alienate children, but rather will offer an interaction interesting and sophisticated enough for adults to enjoy.

IMPLEMENTATION

In order to explore the potential of SSARI through user evaluation, we built an over-life size version of a zebrafish to serve as projection canvas (cf. fig. 2). This zebrafish prototype allows users to explore different layers of content (among other the outer skin, an X-ray visualization and a basic anatomical schematic) by creating shadows and by moving closer to and further away from the fish. While the interaction system could have been built around the exhibition of any 3-dimensional object, we chose to portray the zebrafish due to its relatively simple shape and ongoing research using zebrafish data at our faculty [11, 10].

Figure 2: Prototype zebrafish model.

Based on a previous SAR study involving two projectors [12], we set up two projectors for the display of content. The first projector, referred to as primary projector, was positioned in front of the zebrafish model and was used to project a looping video of the zebrafish skin on the model (cf. fig. 3).

A Kinect motion controller (Microsoft, 2010) was positioned directly on top of the primary projector and used to capture the shadow and the distance of interacting users[1]. The secondary projector was placed on an elevated position and was used to project content on the shadow areas created by the users. In the default state of the prototype, only the primary projector displays content. Once users enter the projection path between the primary projector and the zebrafish model, a shadow is created on the physical canvas. This shadow is then ‘filled in’ with supplementary content by the secondary projector (cf. fig. 4).

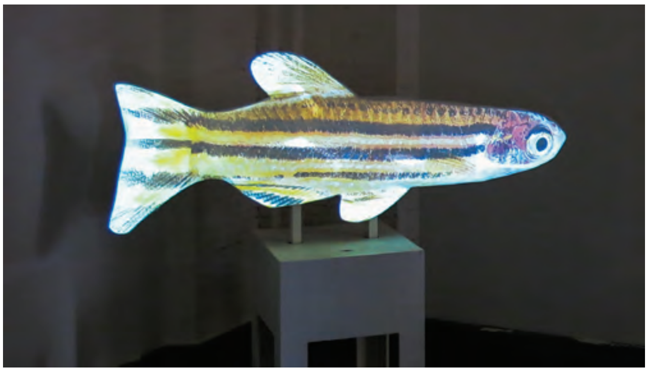

Figure 3: Prototype zebrafish model with skin projection.

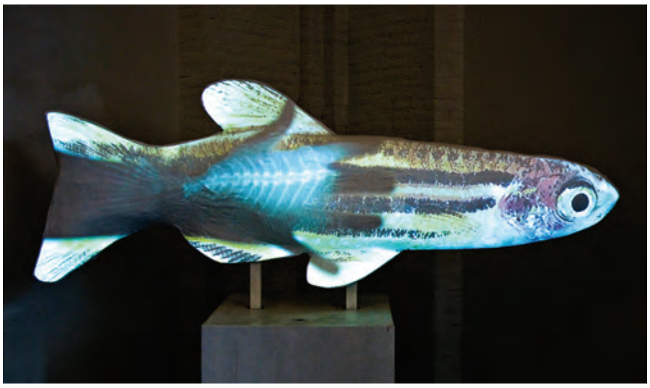

Figure 4: Prototype zebrafish model with supplementary X-ray projection on the shadow area, formed by a hand.

The projection of such content is slightly delayed due to the latency of the Kinect device. As such, fast movements can temporarily cause a visible offset between the position of shadows and the position of the projected supplementary content. It should be noted however that users have not mentioned this delay during our usability assessments (cf. section ‘Usability Assessment’).

The video content was graphically deformed in such a way to accommodate the physical canvas shape and to correct slight distortions in the setup. The video further included small light distortions (caustics) to simulate the effect of the fish being under water. The effect was added in order to draw attention to the display. It also provided a subtle visual atmosphere of the natural environment of the exhibition subject.

At the beginning of our study, we envisioned to use a single layer of supplementary information that users would be able to reveal through creating a shadow on the canvas. However, in response to a first user assessment (cf. section ‘Usability Assessment’), we decided to include multiple layers (cf. fig. 5) for the secondary projection. Users would then be able to explore four supplementary layers: 1) a basic anatomic schematic, 2) a transparent albino specimen, 3) an X-ray visualization, and 4) a histological view.

[1] From a technical perspective it was not the visible shadow that was captured but rather the depth information (e.g. the user’s body) captured through the infrared capabilities of the Kinect. By positioning the primary projector as closely to the Kinect as possible, we were able to align the visible shadow to the captured depth information. A small remaining offset between the two devices was corrected on a software level.

Figure 5: Visualization layers used in the prototype (from top to bottom): outer skin, basic anatomic schematic, transparent albino specimen, X-ray visualization, histological view.

From early user assessments with a simplified prototype (cf. section 'Early Prototype Assessment') we learned that users might need to be actively invited into the interaction area of the primary projector. As a solution we installed a simplified 'zebra crossing' (cf. fig. 6) leading from the primary projector to the fish canvas. Each stripe was placed on the floor and labeled in such a way as to indicate to the user what secondary visualization layer would be revealed if they were to create a shadow when standing at that particular distance. Additionally, we created a stick-figure pictograph sign that was mounted directly under the fish canvas, informing users that interaction would be triggered by stepping into the projection path of the primary projector.

To switch between individual layers, users could move closer or further away from the model. The transition from one layer to another was implemented through a fade over.

On the software side of our prototype, we created a program in 'vvvv' (a node-based visual programming environment) that interprets the measurements of the Kinect depth sensor to control the output of the secondary projector. The program has to be calibrated whenever the positioning of the physical model, the Kinect or the projectors changes in order to correctly project onto the user's shadow. To calibrate the Kinect we used existing code written in 'vvvv' [19] that calculates the required distortion of display content by going through a series of user guided calibration steps. In order to accurately distort the projection content, the program also needs a virtual 3D model of the physical model, which was created by scanning the physical model through handheld operation of the Kinect. We consequently aligned the virtual 3D model with the physical model by adjusting the projection settings. As a result any texture that is applied to the virtual model in our program corresponds to the physical model.

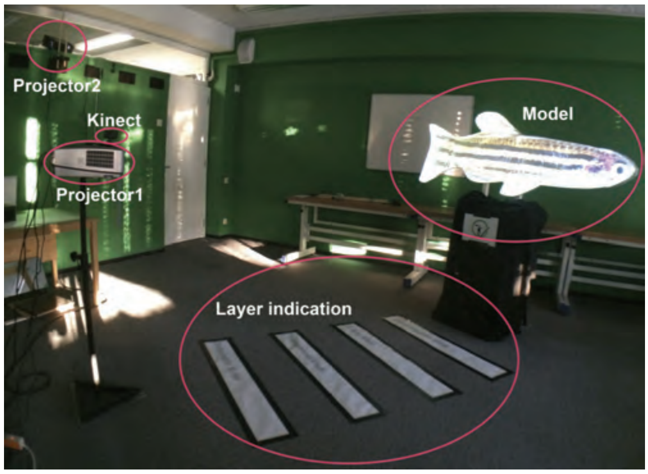

Figure 6: Prototype setup for second user assessment.

To implement the transition of different visualization layers we created an image sequence of the visualizations that served as texture for the virtual 3D model.

The depth values recorded by the Kinect in the case of user interaction are used to select the image frame that corresponds to the measured specific physical depth. Furthermore, we use the depth image of the Kinect to create a matte that blocks out the projection of any content not covered by the shadow of a user. A schematic of the interaction processes with our software can be seen in figure 7. The core usability requirement that our prototype had to fulfill was that its functionality would have to be self explanatory through exploration by users. This meant that once a user had understood that the primary projection path could be interrupted to show supplementary information, all other functionality would be revealed by users through experimenting with the interface.

Finally, we would like to point out that SSARI can generally also be implemented by using only a single projector. In such a case the projector would have to be set up in such a way that interacting users cannot interfere with the projection path. While the projection of the user's shadow would only be simulated, implementing SSARI in this way would reduce the financial requirements. On the other hand, using two projectors allows users to easily understand from where shadows are projected and therefore where they can interact with the display.

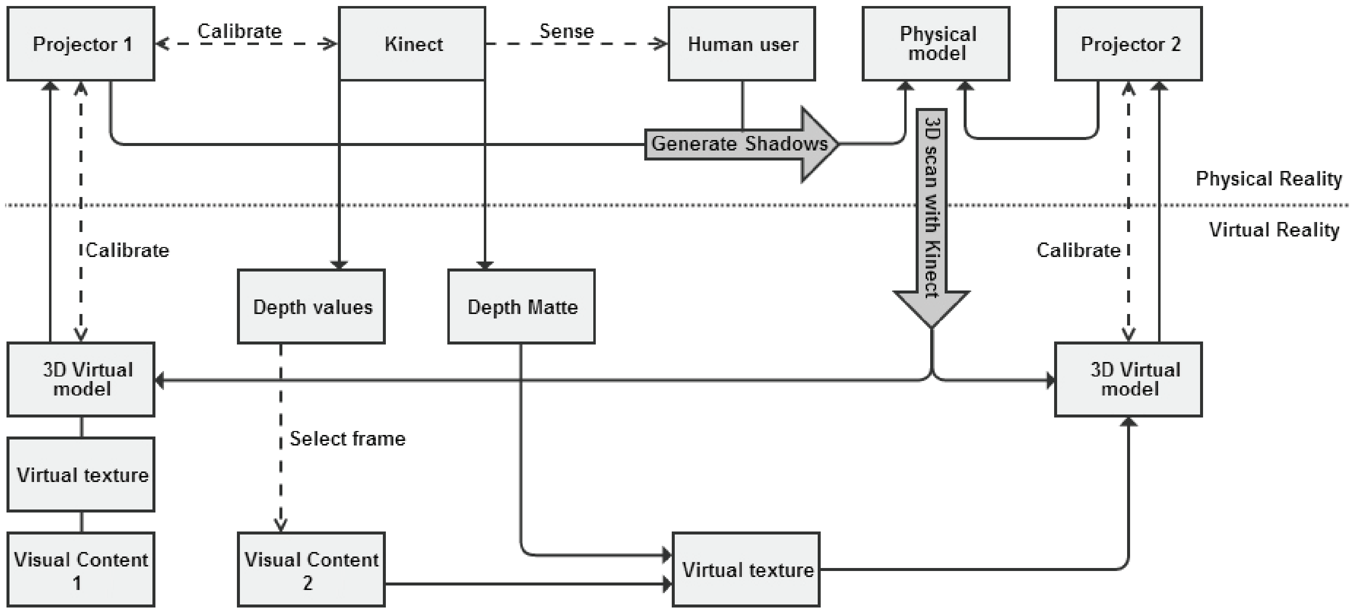

Figure 7: Schematic showing the interaction processes with our prototype software.

Implementation Issues

In the development of our prototype we had to accept some limitations due to constraints in time and resources. One such limitation is the fact that users were still able to interrupt also the secondary projection path if they got too close to the physical model. In such a case, both the primary and the secondary projections would be occluded without means to fill in the shadow area, resulting in a visible (black) shadow. Consequently, our implementation of SSARI does not allow users to touch the model. We discuss possible solutions to this issue later in this paper (cf. section 'Future Work').

Another limitation can be seen in the fact that while multiple users could interact with the prototype at once, the determination of which layer to visualize was dependent on whichever user was closest to the model. As such the layer indications positioned on the ground (in form of a zebra crossing) only corresponded to the position of the user closest to the model.

One constraint that applies when multiple users interact with the system is that our current implementation does not yet display multiple shadow layers at the same time. Consequently, only one user is in control over which content is displayed in all shadow areas.

On a technical level it should be noted that the depth sensing capabilities of the Kinect are low in resolution, resulting in pixelated edges around the detected shadows. Furthermore, fast movements caused a delay between the visual shadow and the supplementary projection.

USABILITY ASSESSMENT

To assess the usability of SSARI we devised two user experiments at different stages of the research. An early prototype was built for the first experiment to explore potential functionality and use cases of SSARI. The second user experiment involved the zebrafish model described previously in the 'Implementation' section.

Early Prototype Assessment

In our first usability experiment we developed a simplified version of the SSARI system. Two projectors were used to project content onto an inoperative computer screen that had been painted white. The primary projector was positioned in front of the computer screen and displayed a looping video of an operating system in use. A secondary projector was positioned at an angle of 60 degrees to display the image of an electronic circuit board. As described previously, a Kinect was used to capture any shadows created by users. This information was then interpreted by our program to display the circuit board imagery on the shadow areas (cf. fig. 8). This early prototype did not change content based on the depth of the shadow. Users could therefore explore two layers at this point: the primary projection and a single shadow-area projection.

We then invited a group of five Master students with backgrounds in interface design to participate in a focus group as expert users. The focus group participants were asked to explore the prototype as a group. During this time we observed the participants and took notes. After approximately 5-10 minutes, the group had decided it had seen enough to start a discussion about possible use cases. Other topics of the discussion were: positive and negative aspects of the interaction, how to invite users to disrupt the primary projection path, and words that described the interaction system. Finally, we asked participants to openly rate the prototype in terms of usability and innovation.

In general the prototype received positive responses in the context of use as exhibition system. Some participants criticized the limited use of interaction possibilities of the early prototype. We were furthermore advised to make sure that future users understand that the projection path can, and in fact should be, interrupted in order to facilitate interaction. During the focus group session, participants seemed to focus on exhibition uses of the interaction system, noting its ability to attract attention and its novel way of interaction. This was also attributed to the fact that the interface did not allow much varied functionality to perform complex tasks. Participants quickly came up with ideas that involved interaction in depth, noting that this dimension could be used to travel back and forth in time or to zoom on a detail. The prototype was rated by each participant, resulting in an average usability rating of 6.8 and an average innovativeness rating of 6.6 (with 1 being worst and 10 being best).

Figure 8: (Left) Painted computer screen used as canvas in the first user assessment, (Right) First user assessment shown in action.

It should be noted that while we were interested in how participants would score the prototype, our primary goal was to spark a conversation between participants that would uncover problematic or ill-defined aspects of the prototype and the SSARI system in general. In many cases, we had already considered specific design solutions mentioned by the participants for implementation in the subsequent prototype. By observing the group discussion we were however able to make an early evaluation regarding what kind of interactivity might be expected from users. Most importantly, it provided an early indication for the suitability of SSARI as form of interaction with exhibition content. As such the focus group session should only be partially understood as usability assessment and also as an early evaluation of the interaction concept itself.

Zebra Fish Display Assessment

To evaluate the usability and user enjoyment of our interactive zebrafish setup, we chose to conduct a comparative experiment. We created a small website for the experiment that contained links to each individual zebrafish visualization. During the experiment, the tablet device was used to connect to the website to simulate a simple touch-based information display. The image resolution of the visualizations was chosen at twice the size at which the image would be displayed on the website. This allowed users to enlarge the view of a visualization using a 'zoom-out' finger gesture while retaining a consistent image quality. When zoomed into an image, participants were able to use finger gestures to pan the image or zoom out again.

To ensure that the comparison between the two systems is as fair as possible, we tried to provide equality in the quality and quantity of interactions available. In addition to the visualizations shown on the SSARI setup, the website therefore also featured a high-resolution image viewer for displaying the histological view. The viewer has been created by the Pennsylvania State University [16] and was linked to in such a way that participants would assume it to be part of the experiment website. This additional visualization was added only to the tablet set up to make up for functionality it did not offer.

The experiment took place with 16 participants between 21 and 38 years of age that had been invited from the faculty premises. To test how the interaction setups would be assessed in a multi-user setting, we asked participants that knew each other to perform the experiment at the same time. A total of 12 participants took the experiment in pairs of two while the remaining 4 took the experiment alone.

Participants were told to take as much time as they wanted to explore the display content after which they would be asked to explore the same content on a different interface. In doing so, participants experienced both setups and could then be asked to compare them. The order in which participants used the two setups was switched after each of the consecutive test sessions. Once participants had interacted with both setups they were asked to fill in a semi-structured questionnaire. Questions included familiarity with touch devices, frequency of museum visits, interestingness of the presented content, ease of use for each of the displays, enjoyment when using each of the displays, and questions regarding the suitability of the setups for use in museums.

RESULTS

The two setups scored roughly the same on average regarding their ease of use, resulting in ratings of 5.9 (touch display) and 5.8 (projection display) out of 7. When asked for their preference between the two for ease of use, exactly half (50%) of the respondents chose the touch display and half chose the projection display.

In terms of self-reported enjoyment, the touch display scored a 4 out of 7 on average, while the projection display a 6.4 out of 7. Asked to choose which of the two displays was more fun the majority of the respondents (93.8%) preferred the projection display. When inquiring about the suitability in museums, the touch display scored 4.2 out 7 on average. Here the projection display scored an average of 6.4 out of 7 with no rating being lower than a 5. In terms of how informative the two displays were perceived relative to each other, 7 respondents (44%) found the projection display more informative, while 9 respondents found the touch screen more so.

When asked via an open question for any further comments or opinions they might have, respondents indicated that they would like additional information to accompany the content either through text or voice narration. The ability of the projection display to accommodate multiuser interaction on its own was described as an advantage over touchscreen displays by two of the respondents. On the other hand, quality and detail of information displayed by the projected display was criticized as poorer relative to the touch screen display. Regarding our usability requirements, all participants discovered the four supplementary information layers without being prompted to explore them.

CONCLUSIONS AND DISCUSSION

The results of our user assessment experiment suggest that the majority of users found the SSARI system suitable for use in museums, and significantly ($p = 0.0001$) more suited than the use of the touch display (H1 and H5 confirmed). In terms of ease of use, users found our prototype somewhat easy to use, rating it with an average of 5.8 out of 7. The touch display was rated an average of 5.9 and therefore slightly easier to use. While the difference is not significant ($p = 0.42$), our second hypothesis cannot be confirmed (H2 rejected). In regards to enjoyment of the interaction, users reported consistently high enjoyment (6.4 out of 7) of our SSARI implementation, with 15 out of 16 participants finding it more enjoyable than the touch display (H3 confirmed). Lastly, the majority of users (9 out of 16) stated that the touch display was more informative (H4 rejected).

Overall we conclude that the concept of SSARI is practically feasible, and (from a user perspective) suited to be used in museum environments. Participants in our research were enthusiastic about exploring the display content through SSARI. Our experiments show that the strengths or SSARI are found in how content is displayed and interacted with. Given the fact that the tablet system was found to be more informative yet less suited for museum use, we conclude that users may prefer traditional forms of displays when trying to access more in-depth information. As such we see the ideal use of SSARI in a complementary setting where detailed information is made available through other means (e.g. textual or auditive).

A SSARI setup that includes depth-sensing of interacting users is particularly suitable for exploring different layers of content that correspond to the same physical shape. Another strength of SSARI is that it allows several users to interact with the display at the same time and hence also allows for interaction and exchange among different users. In our prototype, users shared their shadows in such a way that the same depth visualization layer was shown for each of them. However, the system could also be set up so to give each user full depth control, and therefore selection of layered display content, of his or her own shadows.

As we pointed out earlier, SSARI can technically be implemented with a single projection. However, we argue that doing so also reduces the immediacy and playfulness of the setup. In the end, we do not consider SSARI a technique that excels in productivity but rather in user enjoyment. Efforts regarding a simpler implementation should therefore be mindful to not lose the appeal that it provides to its users.

Finally, in addition to being used in museum environments, we see interesting potential for SSARI in artistic performances and installations, as well as the implementation in experimental games and educational setups. It could, for example, be used to learn about human anatomy at schools. We hope that our research can inspire the creation of more examples involving SSARI.

FUTURE WORK

In the future we will further experiment with different forms of interactions and visualizations in our zebrafish prototype. One possibility we are currently exploring is the visualization of histological views in such a way that users can interactively 'slice' through the zebrafish by moving closer or further away. This functionality has been added to the prototype setup but has not yet been evaluated with user experiments.

As mentioned in the previous section, a future step in terms of interaction would be to implement the possibility to display multiple shadow layers at the same time. This way, each user would always be in control over which content to display regardless of the interactions of other users. Furthermore, we are interested in adding a secondary, traditional display to the prototype for complementary use. Interacting with the fish model could for example trigger the display of more in-depth information on a secondary screen (cf. fig. 9).

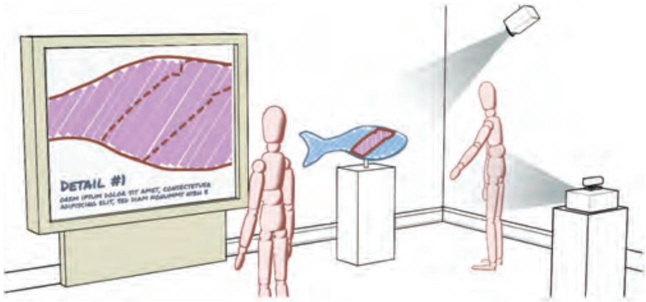

Figure 9: Schematic showing the zebrafish prototype complemented by a secondary screen, showing a zoomed-in version of the secondary display layer.

With respect to the implementation, we want to create a more solid installation. From a technical standpoint, we would like to perfect the match between the physical model and the virtual projection. This could for example be achieved by 3D printing the physical model based on the particular virtual model that will be projected onto it. We also consider recreating the prototype in such a way that the second projection would occur from within or behind the model (given the use of semi-transparent building materials). Modifying the prototype in such a way would allow users to come very close to the model and even touch it; enabling even more possibilities for interaction.

In regards to our research, we recognize that our installation may have scored high on enjoyment due to the novelty of the interaction as well as the novelty in the presentation of content. While both aspects have been implemented in other projects before, it is unlikely that many users would consider them common. Future research should therefore evaluate whether the use of SSARI systems provide consistently high user enjoyment, and if so, what the underlying reasons may be. In the end it may very well be that the success of SSARI depends on its ability to surprise users.

Another aspect to consider is the fact that this study has focused on technical feasibility of SSARI on the one side and its user reception on the other. The suitability of a technology however also depends on the involvement of other stakeholders. When considering the use of SSARI in museum environments, future research of SSARI should investigate aspects such as technical maintenance and upgradeability.

We hope that this research inspires museums and artists alike to use SSARI as well as participate in its ongoing research.

ACKNOWLEDGEMENTS

We would like to thank all the participants that took part in our user experiments. Furthermore we want to specifically acknowledge Meggy Pepelanova, Georgios Lampropoulos, Livia Teernstra, Isabelle Kniestedt, and Bram Snijders for their support in this research.

REFERENCES

1. Apperley, M., McLeod, L., Masoodian, M., Paine, L., Phillips, M., Rogers, B., and Thomson, K. Use of Video Shadow for Small Group Interaction Awareness on a Large Interactive Display Surface. 81–90.
2. Bimber, O., and Raskar, R. *Spatial augmented reality*. Merging real and virtual worlds. A K Peters Limited, 2005.
3. Deterding, S., Dixon, D., Khaled, R., and Nacke, L. From Game Design Elements to Gamefulness: Defining Gamification. 9–15.
4. Herndon, K. P., Zeleznik, R. C., Robbins, D. C., Conner, D. B., Snibbe, S. S., and Van Dam, A. Interactive Shadows. 1–6.
5. Hornecker, E., and Stifter, M. Learning From Interactive Museum Installations About Interaction Design for Public Settings. 135–142.
6. Jacquemin, C., Gagner´e, G., and Lahoz, B. Shedding Light on Shadow: Real-Time Interactive Artworks Based on Cast Shadows or Silhouettes. 173–182.
7. Kidd, J., Ntalla, I., and Lyons, W. Multi-Touch Interfaces in Museum Spaces: Reporting Preliminary Findings on the Nature of Interaction. *Rethinking Technology in Museums: Emerging Experiences. University of Limerick* (2011).
8. Maynes-Aminzade, D., Pausch, R., and Seitz, S. Techniques for Interactive Audience Participation. In *ICMI '02: Proceedings of the 4th IEEE International Conference on Multimodal Interfaces*, IEEE Computer Society (Oct. 2002).
9. Miwa, Y., and Ishibiki, C. Shadow Communication: System for Embodied Interaction with Remote Partners. 467–476.
10. Potikanond, D., and Verbeek, F. Visualization and analysis of 3D gene expression patterns in zebrafish using web services. In *IS&T/SPIE Electronic Imaging*, International Society for Optics and Photonics (2012), 829412–829412.
11. Potikanond, D., Verbeek, F., et al. 3D visual integration of spatio-temporal gene expression patterns on digital atlas of zebrafish embryo using web service. *ACEEE International Journal on Information Technology 1*, 3 (2011).
12. Raskar, R., Welch, G., Low, K.-L., and Bandyopadhyay, D. Shader Lamps: Animating Real Objects with Image-Based Illumination. 89–102.
13. Samosky, J. T., Nelson, D. A., Wang, B., Bregman, R., Hosmer, A., Mikulis, B., and Weaver, R. BodyExplorerAR: Enhancing a Mannequin Medical Simulator with Sensing and Projective Augmented Reality for Exploring Dynamic Anatomy and Physiology. 263–270.
14. Shneiderman, B. Touch Screens Now Offer Compelling Uses. *Software, IEEE 8*, 2 (1991), 93–94.
15. Shoemaker, G., Tang, A., and Booth, K. S. Shadow Reaching: a New Perspective on Interaction for Large Displays. In *UIST '07: Proceedings of the 20th annual ACM symposium on User interface software and technology*, ACM Request Permissions (Oct. 2007).
16. The Pennsylvania State University. Zebrafish Atlas VirtualSlide Viewer, 2013.
17. Verlinden, J. C., De Smit, A., Peeters, A. W., and van Gelderen, M. H. Development of a flexible augmented prototyping system. In *WSCG* (2003).
18. vom Lehn, D., and Heath, C. Accounting for New Technology in Museum Exhibitions. *International Journal of Arts Management* (2005), 11–21.
19. Woods, E., and McDonald, K. Notes from Art&&Code: Calibrating Projectors and Cameras: Practical Tools, 2011.
20. Xu, H., Iwai, D., Hiura, S., and Sato, K. User Interface by Virtual Shadow Projection. 4814–4817.

Chi sparks 2014

Auditory Signal Design for Automatic Number Plate Recognition System

C.G. Heydra
Faculty of Communication & Multimedia Design, The Hague University of Applied Sciences
Stamkartplein 40
2521 EP Den Haag
C.G.Heydra@hhs.nl

R.J. Jansen
Faculty of Industrial Design Engineering, Delft University of Technology
Landbergstraat 15
2628 CE Delft
R.J.Jansen@tudelft.nl

R. van Egmond
Faculty of Industrial Design Engineering, Delft University of Technology
Landbergstraat 15
2628 CE Delft
R.vanEgmond@tudelft.nl

ABSTRACT
This paper focuses on the design of an auditory signal for the Automatic Number Plate Recognition system of Dutch national police. The auditory signal is designed to alert police officers of suspicious cars in their proximity, communicating priority level and location of the suspicious car and taking into account the auditory environment of the police car. Design goals are formulated and corresponding design principles are applied and tested. Conclusions are drawn and discussed and recommendations for future work are made.

Author Keywords
Audio design; Auditory signal design; Localization, Automatic number plate recognition plate system; Dutch national police

ACM Classification Keywords
H.5.2 User Interfaces > Auditory (non-speech) feedback

H.5.5 Sound and Music Computing > Signal analysis, synthesis, and processing

INTRODUCTION
The Dutch national police are currently experimenting with a system to aid identifying suspicious cars, in order to apprehend criminals or other persons of interest. This Automatic Number Plate Recognition (ANPR) system uses two roof-mounted cameras to read the licence plates of all cars in scanning range of a police car. Subsequently, the licence plate numbers are compared to a police database, to determine whether the car or its registered owner is of interest for the police. In that case, the police officers receive visual and auditory feedback.

There are several causes why a car can be of interest for the police, e.g., a stolen vehicle, or a suspended license plate. The former case is of higher priority than the latter case. Internal communication with the Dutch national police resulted in a selection of four priority levels for ANPR signals.

The ANPR system interrupts officers during their current task. According to priority level, the officers themselves decide whether to respond to this suspicious car or carry on with their current activity. In order to make this decision, the signal of the ANPR system should be intuitive, concise and informative.

This paper focuses on the design of an auditory signal of the ANPR system that is effective in the auditory environment of the police officer.

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

PROBLEM SITUATION
In the current system, police officers always receive the same alarm sound. For several reasons, this is not an optimal solution. First, the alarm sound does not differentiate between causes with a high or low priority. Second, the current activity of police officers is not taken into account. The alarm sound does not differentiate between causes with a higher or lower priority compared to the current activity. Consequently, police officers always have to examine the screen to decide on whether to take action (e.g., [4]). Third, the alarm sound does not differentiate between parked and moving cars. In the latter case, police officers have the opportunity to take action immediately. Fourth and finally, the alarm sound does not convey a sense of direction. Two frontal loudspeakers produce the monophonic alarm sound. Thus, the centre of the car is perceived as spatial location of the sound. In case of multiple lanes of traffic, police officers first have to examine the dashboard monitor to know on which side of the car they should look. This process often takes too long to be able to take immediate action.

THE AUDITORY ENVIRONMENT
In addition to visually monitoring the environment, police officers continuously monitor incoming radio messages. On the one hand, they listen if a message is directed at them (i.e., being assigned to a call). On the other hand, they listen in on messages intended for colleagues to stay informed on their whereabouts and duties. The amount of messages may vary from once every 10 minutes to 30 messages in 5

minutes. Typically, the driver wears an earpiece in his/her left ear. In case of duo surveillance, the co-driver wears an earpiece in his/her right ear. Additionally, messages can be played through the car's loudspeakers.

At high driving speeds there are two main sources of car sounds: rumbling engine sounds, and wind sounds caused by the non-aerodynamic light bar. In case of emergency, the combination of a loud siren with the car sounds makes it nearly impossible for police officers to talk with each other. Consequently, police officers crank the volume of their earpieces, which in turn leads to complaints about fatigue in the long run.

When analysing the engine and wind sound, it can be concluded that these sounds are most intense below 200Hz. Wind noise is also present in the range between 200Hz and 7000Hz, but in much lower intensity.

DESIGN GOALS

The sounds designed for the ANPR-system should be intuitive, concise and informative in the auditory environment of the police officer. Therefore, the following five design goals are formulated.

1. The signal should be distinctly audible in the auditory environment of the police car.
2. The signal should convey the priority level of the suspicious car.
3. The signal should convey the location of the suspicious car in respect to the police car.
4. The signal should convey the direction of movement of the suspicious car in respect to the police car.
5. The signal should distinguish stationary cars from moving cars.

USERS

Although based on generalisations, some common characteristic can be attributed to the personality of police officers, such as acceptance of hierarchy and strong leadership, a preference for conciseness, and a need for clarity in procedures and protocol [1, 2]. These characteristics can be used to determine the main concepts the signal should convey: dependable, commanding, concise and assertive. In turn, these main concepts can be translated to acoustical properties, as shown in Table 1.

Main concept	Acoustical properties
Commanding, concise, assertive	Staccato, short attack, short decay
Dependable, assertive	Muffled or low pitched
Commanding, assertive	Some grittiness
Commanding, concise, assertive	Short sounds

Table 1: acoustical properties and corresponding main concepts

DESIGN

Selecting frequency range

As mentioned earlier, the auditory environment includes engine and wind noise, which are most intense below 200Hz. Therefore, any designed sound should have a frequency higher then 200Hz for better audibility. Selecting a certain frequency range also affects the user's ability to pinpoint the location of where the sound originated. The accuracy of localization is dependent on frequency and angle of the sound [6]. Full frontal sounds have a broader localizable frequency range than sounds at an angle. As Figure 1 illustrates, the frequency range at which a sound is localizable for the most angles are the frequencies below 1000Hz and between 2800Hz and 4000Hz. Considering the requirements for the auditory environment (design goal 1) and localization (design goals 3 & 4), the designed signal should predominantly use frequencies in the ranges of 200Hz to 1000Hz and 2800Hz to 4000Hz.

Message composition

The message conveyed by the designed signal is composed of three elements, which communicate the following message: *"This is the ANPR-system, there is a suspicious car over there"*. These elements, namely identification (design goal 1), priority level (design goal 2), and location (design goals 3 & 4), will be discussed next.

Identifying the ANPR-system

Identification of (and a call for attention to) the ANPR-system aims to provide a frame of reference for the police officers. In this case, a header sound is designed. This header is a sound, prior to the sound notifying priority and location. Its goal is to prepare the users for a confined number of possible notifications. Therefore, it directs the user's attention to the upcoming task of choosing whether to respond or ignore the upcoming notification of a suspicious car.

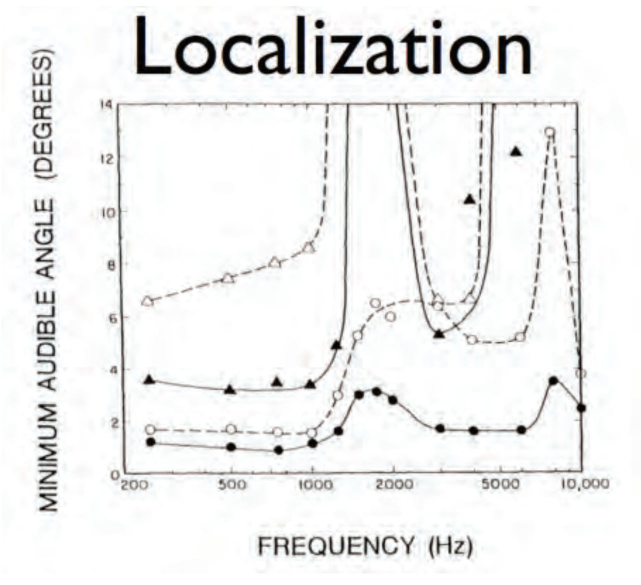

Figure 1: The minimum audible angle between successive pulses of a tone as a function of the frequency of the tone and the azimuth of the source (●=0°; ○=30°; ▲=60°; △=75°). Adapted from [6]

The design of the ANPR header sound is based on the Dutch pronunciation of the abbreviation ANPR and therefore consists of four separate phonemes. As the waveform in Figure 2 shows the letters A and N blend into each other, where the letters P and R are more self-contained.

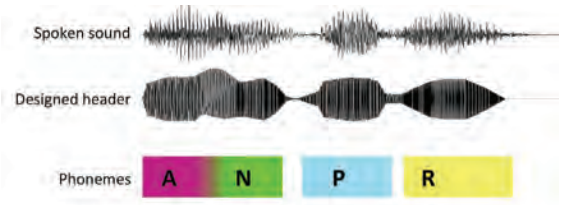

Figure 2: Header waveform and phonemes

The designed header coheres to the iamb of the spoken sound, but it does not cohere to the pitch because of the frequency space covered earlier. The total duration of the header is 400ms instead of the 1000ms the spoken sound. This duration is made shorter, because the speed at which the car travels requires a fast response time. The sound is composed of simple sine waves with a base frequency of 300Hz and its 2nd, 3rd and 4th harmonic.

Priority level

After the header sound, the signal continues by communicating the priority level. This priority level sound consists of a combination of two 200ms long sounds, one with a base frequency of 300Hz, and one with a base frequency of 220Hz. Harmonics are used to increase pitch robustness [5], and to reduce potential masking effects by other sound sources. Taking into account the frequency range for optimal localization described earlier, both signals contain the 2nd, 3rd, 9th, and 10th harmonics. These two sounds are used in conjunction with the following two principles to convey the four priority levels.

1) Texture of the signal. Two of the main concepts that appeal to the users are assertive and commanding. This is translated to the texture of the signal. A more gritty texture, meaning short, jagged fluctuations in the volume of the harmonics, of the signal will be perceived as a more assertive and commanding signal. Therefore as the priority of the signal increases the grittiness also increases.

This alone cannot be enough to distinguish the priority levels because it works by comparison. When two sounds are played one after another, it is easy to say which is grittier, but without a comparative context, other cues must be implemented to obtain a clear distinction.

2) Meter of the signal. The meter refers to the number of beats in a bar of music, and determines how a repetitive pattern (rhythm) occurs. Often, but not always, a rhythm with a 2/4 meter is experienced as more lively, erratic and jittery then a 4/4 beat. Therefore the chosen meters for the priorities 4, 3, 2 and 1 are 4/4, 3/4, 2/4, and 1/4, respectively. By combining the high and low pitched sounds and the time signature, the final priority level sounds were created as visualised in Figure 3.

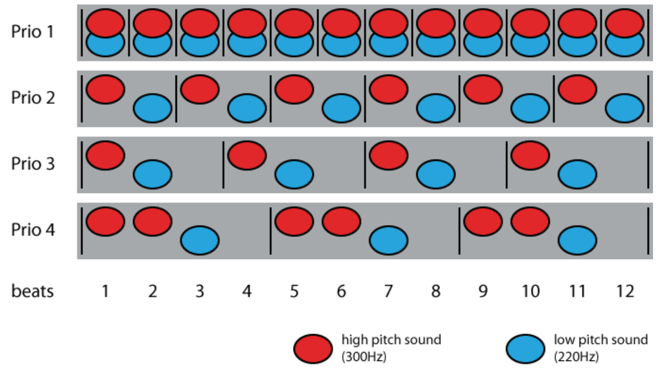

Figure 3: sequences of sounds for each priority level sound.

Localization

As specified in the design goals the location and direction of movement of the suspicious car should be communicated by the signal (design goals 3 & 4). Three principles are used to address this localization and spatiality in accordance with the aforementioned frequency range.

1) Sound origin. Installing multiple speakers inside the police car provides an opportunity for spatiality, by differing the volume (and applying subtle timing changes) between the speakers. This gives the impression of sound coming from the left, right, front, back or anywhere in between. Alternatively, spatiality may be obtained through binaural techniques by mounting two speakers at each seat's headrest (e.g., [3]).

2) Volume. Changes in sound volume can be used to convey changes in distance between a listener and a moving source [9]. In the current design, a lower volume indicates that the suspicious car is further away. This volume drop is only applied behind the car, because detecting a car at great distance in front of the police car then would lead to a signal with a volume too low to determine the priority (or even too low to notice). Figure 4 shows the location of the volume drop.

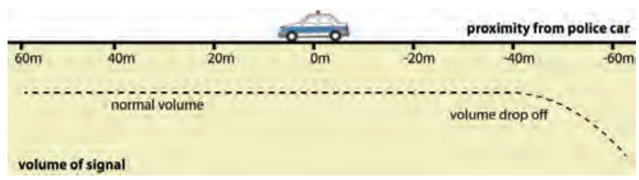

Figure 4: Location of volume drop off.

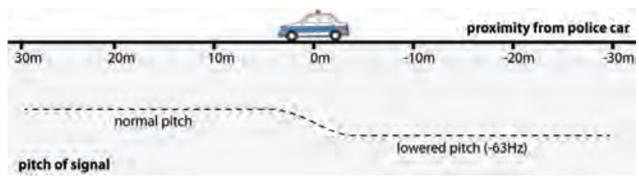

Figure 5: Location of pitch change.

3) Doppler like effect. The third localization cue is a Doppler like effect that occurs when a suspicious car passes the police car (or vice versa). If the suspicious car is located behind the police car the pitch of the signal will be noticeable lower (63Hz) then when it is in front of the police car. If the two cars pass each other a Doppler effect is quite noticeable in the signal. Figure 5 shows the location of this pitch change in respect to the police car.

Stationary versus moving cars

It is beneficial for a police officer to distinguish moving cars from stationary cars in order to quickly identify the suspicious car. Therefore the signal for stationary cars differs from the sound for moving cars (design goal 5). A low tone (base frequency 100Hz and 2^{nd}, 3^{rd}, 9^{th} and 10^{th} harmonic, duration 600ms) is chosen to signal the officer that the car is stationary. Although the base frequency is outside the preferred frequency range, the lower and longer tone is chosen to evoke a sense of mass and immovability. When a stationary suspicious car is detected, the signal starts the same as that of a moving car, i.e., a header followed by a priority level sound. But at the point the suspicious car is next to the police car, the low tone sound is played either at the left or the right speakers. This indicates that the suspicious car is parked (or otherwise stationary) to the direct left or right of the police car.

TESTING

The designed signal is tested in three different ways to measure the recognizability and learnability of the priority level sound and effectiveness of the localization.

Test 1: recognizability of priority levels

A first-time user should be able to determine the priority level of the four sounds. Therefore, this test aims to answer the question: *Is the priority level intuitively recognizable?* Five participants used headphones to listen to the priority level sounds in random order, and where given the task to arrange the four sounds in order of perceived priority. They were allowed to listen to the sounds again if they wanted to. The results of this test (see Table 2) show that none of the participants perceived all priority levels as they where designed. Prio1 and Prio2 are swapped in half of the results. This makes them intuitively ineffective. Prio3 and Prio4 are swapped in all but one case. This could make them intuitively effective, but just not in the way they were designed. Furthermore, the results show no confusion between Prio1 & Prio2 on the one hand, and Prio3 & Prio4 on the other hand. Overall, this test shows that the two principles used in designing the priority level sounds do not seem to work as intended.

Test 2: learnability of priority level sounds

This test addresses the question: *Once primed with the reference sounds, are the priority level sounds distinguishable by the participants?* In order to test this, five other participants used headphones to listen to the four priority sounds, and were told to which priority they belonged as a reference. They listened to these four sounds twice in random order. Participants were instructed to identify each of these eight sounds. Participants were not allowed to replay the sounds, nor were they allowed to revisit their previous answers. The results are shown in Table 3.

In this test 36 of 40 priority level sounds were identified correctly over all participants after they were primed with the reference sounds. So, it can be concluded that in 90% of the cases, the priority level sounds are correctly distinguished once they are familiar with the sounds.

Participant	Prio1	Prio2	Prio3	Prio4
A	2	1	4	3
B	1	2	4	3
C	2	1	4	3
D	1	2	4	3
E	2	1	3	4

Table 2: Results recognizability test. Designed priority order versus intuitively perceived priority order.

Participant	Correctly identified	Incorrectly identified
F	5	3
G	8	0
H	8	0
I	8	0
J	7	1

Table 3: Results learnability test.

Scenario	Description
1	Car moving in the same direction, appearing in front right of the police car. The car is overtaken by the police car and finally appears alongside the police car on the right.
2	Car moving in the oncoming direction. Appears from a side street in front left of the police car.
3	Car parked in a row of cars on the right side of the police car.
4	Car moving in the same direction. Changing lanes and appearing directly in front of the police car.

Table 4: Test scenarios for testing effectiveness of localization.

Participant	Sce1	Sce2	Sce3	Sce4
A	-1	v	x	v
B	v	-1	x	x
C	v	v	-1	v
D	v	v	x	v
E	v	-1	-1	v

Table 5: Results localization scenario test. (v = correctly identified; -1 = identified an adjacent car; x = incorrectly identified)

Test 3: effectiveness of localization

This test aims to answer the question: *Are the designed localization cues effective in identifying suspicious cars?* Four scenarios with suspicious cars were designed and represented as short film clips, see Table 4. Each film clip featured the same footage, including environmental sound, of a driving car filmed from the point of view of the driver. Signal sounds with localization cues were added to each clip.

Five participants, the same as in test 1, were instructed that a sound would indicate a specific moving or stationary car. Their task was to point at this car on the screen. Participants were allowed to watch the video clip again if desired. The results are shown in Table 5. As the results show, 11 of 20 cases were identified successful, and 5 cases where off by one car. Thus, in 55% of the cases the localization was accurate, and in 80% of the cases it was accurate within a margin of one car.

A notable result can be seen in scenario 3, where none of the participants identified the car correctly. Two of the participants identified the parked car in front of the car that was pointed at; the other three participants identified moving cars.

CONCLUSION[1]

To achieve the design goals, the final design of the signal sound for the ANPR-system consists of a header sound and a priority level sound. Signal localization cues are applied to draw attention to the correct position.

Test 1 shows that the priority level of the signal is not intuitively recognizable. An explanation for this could be that the meter was not perceived in the intended way. One thing that was not taken into account in the design was the perceptual rhythmic and melodic accents that can be heard in sequences of sounds. Povel & Essens [7] (rhythmic) and Van Egmond et al.[8] (melodic) show that in a grouping of three or more tones, the first and last tone receive an accent. This could explain the ineffectiveness of priority level 4 sound because it may have been perceived as a 2/4 instead of a 4/4 meter, as illustrated in figure 6. A similar perceptual issue occurs in the priority level 1 sound. Because there are no accents, no meter is evoked.

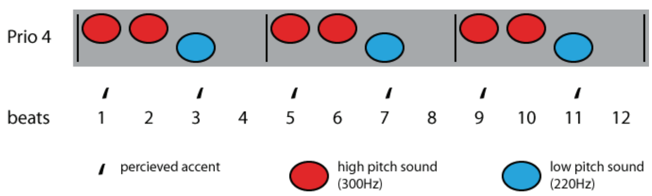

Figure 6: perceived accents in priority level sound 4.

Although the priority level sounds are ineffective in intuitively conveying their priority, test 2 shows that they are clearly distinguishable once the user is familiar with the priority level sounds. Next, test 3 shows that the localization principles applied to the signal were successful in drawing the attention to a particular car. Therefore, this design seems an effective starting point for further research, and to add a new functionality to the ANPR-system.

[1] The second and the third author were involved in the study after the experiments were conducted, in order to give the paper a theoretical framing and for the interpretation of the results.

FUTURE WORK

A first point of improvement would be a better design of the priority level sounds in which the rhythmic and melodic accents are properly applied. This will hopefully increase the intuitive recognizability and help officers in determining the priority level.

Although this study focused on the design of an auditory signal for the ANPR-system, it would be beneficial to broaden the scope. The designed sound should be seen as one element of the interface that consists of more elements that work together in achieving the same goal. Visual devices like dashboard monitors or even integrating visuals in the windscreen may increase the efficiency of the ANPR-system. In this study we employed the method of scenario based testing, using video clips. A next step would be to test the sound design in a natural setting.

REFERENCES

1. Abrahamsen, S. Police Personality: And the Relationship between Personality and Preferences for Conflict Resolution Tactics. M.Sc. thesis, Politihøgskolen, Oslo (2006).
2. Balch, R.W. The Police Personality: Fact or Fiction? *The Journal of Criminal Law, Criminology, and Police Science*, Northwestern University (1972), 63:1, 106-119.
3. Blauert, J. & Rabenstein, R. Loudspeaker Methods for Surround Sound. *Proc. 57th Open Seminar on Acoustics*, OSA (2010).
4. Jansen, R.J., Van Egmond, R., De Ridder, H. & Silvester, S. Transitional Journey Maps: Capturing the Dynamics of Operational Policing. *Proc. Human Factors and Ergonomics Society Chapter Europe Ann. Conf. 2013*, HFES-Europe (2013).
5. Meddis, R. & Hewitt, M.J. Virtual Pitch and Phase Sensitivity of a Computer Model of the Auditory Periphery. I: Pitch Identification. *J. Acoust. Soc. Am.*, Acoustical Society of America (1991), 89:6, 2866-2882.
6. Mills, A.W. & Tobias, J.V. *Foundations of Modern Auditory Theory*. Academic Press, New York, 1972, 303-348.
7. Povel, D.J. & Essens, P. Perception of temporal patterns. *Music Perception: An Interdisciplinary Journal*, University of California Press (1985), 2:4, 411-440.
8. Van Egmond, R., Meulenbroek, R.G.J. & Franssen, P. The influence of perceptual stability of auditory percepts on motor behavior. In De Waard, D., Brookhuis, K.A., Van Egmond, R. & Boersema, Th. (Eds.) *Human Factors in Design, Safety, and Management*. Shaker Publishing, Maastricht, The Netherlands, 2005, 235-246.
9. Zahorik, P., Brungart, D.S. & Bronkhorst, A.W. Auditory Distance Perception in Humans: A Summary of Past and Present Research. *Acta Acustica united with Acustica*, Hirzel Verlag (2005), 91, 409-420.

Presence Mediator: Creating Social Presence through Spatial Positioning of Everyday Objects

Daniel Egerer
Mads Clausen Institute
University of Southern Denmark
Alsion 2, 6400 Sønderborg, Denmark
daege12@student.sdu.dk

Jelle van Dijk
Technology and Innovation
Utrecht University of Applied Sciences
P.O.box 182, 3500 AD, Utrecht, Netherlands
jelle.vandijk@hu.nl

ABSTRACT

In the situation where people live together in one household, the rearrangement of everyday objects (either accidental or deliberate) becomes a trace of social presence that people may use as part of their social interactions. This research aims at simulating that natural phenomenon in the context of a long-distance relationship. That is, mediation of social presence through domestic objects is extended over time and space. Our concept, the Presence Mediator, is an interactive system that is attached to an ordinary domestic object. Participants of a field test were in the belief that two mediating objects are synchronized and spatially representing each other's action in the social relation. In fact, the prototype moved a few times per day at random distance and direction. The results illustrate the complexity mediating objects entail as they bring further consequences for the social relation into play. We discuss the difference between traces of social presence that form part of the common background of attention, versus social interaction as mediated by objects that form the center of attention. Furthermore, we discuss the object as a mediating trace of another person's presence, versus the object as an explicit (anthropomorphic) representation of that other person.

Author Keywords

Tangible user interface; interaction design; social presence; long-distance relationship; embodied cognition

ACM Classification Keywords

H5.2. User Interfaces, H.5.3. Collaborative Computing

INTRODUCTION

Tangible User Interfaces have gained increasing attention within the field of Human-Computer-Interaction in recent years. Tangible Interaction describes concepts that rely on embodied interaction, tangible manipulation, physical representation of data, and embeddedness in physical space [13]. While traditional screen-based solutions enable users to interact with bare virtual information, tangible interfaces incorporate embodied interaction within the concrete world of physical objects and (other) people. Designing tangible interfaces requires consequently not only a sound understanding of the digital, but also of the concrete world that is inhabited by our body, and in particular the relation between the digital and the "embodied" environment within hybrid systems [22].

Related work on Tangible Interaction has shown various definitions and dominant approaches within different disciplines: Within the field of Computer Science and HCI, the focus lies in the coupling of digital data with the physical representation and manipulation [8]. The product design community, however, is concerned with the domain-tailored design of expressive interaction itself and focuses on the bodily interaction, making use of sensory richness and action potential of physical objects so that meaning is created in the interaction [7, p. 288]. Interactive arts and architecture on the other hand deals with physical interactive installations in large-scale spaces that require full-body interaction [13].

Physical objects embody many different qualities interesting from an interaction point of view as it provides sensory richness for the user. Current research focuses predominantly on tangible user interfaces as a means for embodied input to manipulate digital data. This work however goes beyond the goal of designing a representation of digital data in tangible form [15]. We examine whether and how everyday objects, with both their physical and digital properties taken as a whole, can help to create experiences of social presence between people. On a theoretical level, this study is a research-through-design effort exploring the notion of Socio-Sensorimotor Couplings [22]. A socio-sensorimotor coupling is a theoretical notion that combines social interaction and the physical interaction between body and the local environment. The idea, based on principles of Embodied Cognition [4, 14] and Situated Cognition [21, 3] is that people generally make sense of the world while dealing with objects in the environment, both as part of a social coordination between people, and as part of sustained sensorimotor loops that link perception to action [12]. Together, this process of socio-sensorimotor coupling helps a person to make sense, and act successfully in the everyday world [22]. In what follows, we apply this notion to the concrete case of long distance relationships and in reflection on the design case and the user study we present some insights relevant to tangible interaction design for social interaction.

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

TRACES OF SOCIAL PRESENCE

Everyday objects and the embedding environment play a crucial role on how social interaction takes place [17]. A cohabiting couple in an apartment, for example, rearranges its environment continuously. Each action of an individual has an impact on both and leads to new understanding in how they relate to each other. This ongoing transformation is essential part of living and being together. Everyday objects become thus footprints or traces of existence, action and eventually presence in social interaction [22]. The question is whether we can create digitally augmented objects to create social presence at a long distance scale.

Let's consider following scenario based on the notion of traces mentioned earlier: It is Monday morning in an apartment of a cohabiting couple. While one of them (A) is getting up quite early to go to work, the other one (B) is still sleeping. A is taking a shower, brushing his teeth, making coffee, preparing dishes, having breakfast and eventually leaving the apartment. After some time B is getting up and able to speculate what has happened this morning: The shower stall is wet, the toothpaste is sticking on the sink, the coffee can is on the counter and so on. This ordinary everyday situation illustrates that meaning in social presence is not only constituted through explicit forms of human-human interaction, but also by the implicit consequences of human action that are left as traces distributed in the environment [22].

In this work we investigated how this notion can be applied to long-distance relationships by using everyday objects as mediators for social presence. In contrast to the previous scenario, let's consider a couple that is living geographically-separated in two different households. Caused traces by A cannot be perceived by B so that presence is reduced predominantly to its common explicit forms such as phone calls or video chats. In order to overcome this discrepancy, we have developed an interactive system, the Presence Mediator, that is embedded in a domestic object. The Presence Mediator is situated in each household and both are spatially synchronized by sending and receiving each other's position. A is able to reposition the own mediating object that again gets spatially represented at B at the same time and vice versa (Figure 1).

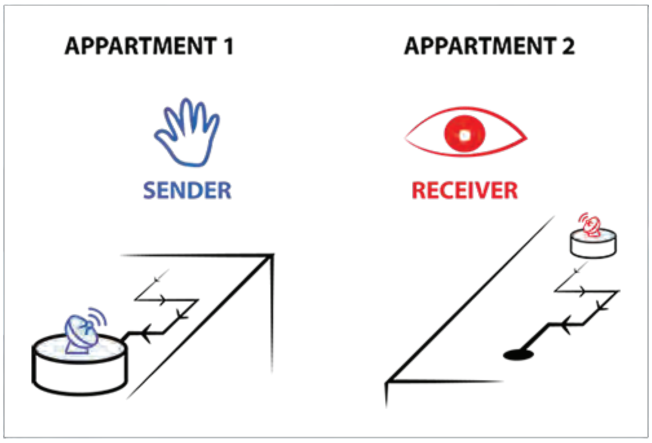

Figure 1. Both objects are bi-directionally synchronized and create similar traces over distance.

Consideration should be given to the ambiguous interpretation of repositioned objects. The objects are intended to convey ambiguity in meaning and open-ended situations over defined utilitarian purposes. This has been investigated with The History Tablecloth, a flexible, screen-printed electroluminescent material [11]. When objects are left on the tablecloth, cells beneath light up that grow over a period of time. The History Tablecloth (Figure 2) elicited rich interpretations among the inhabitants so that local knowledge was co-constructed. Even occasionally erroneous behavior of it led to greater richness in experience, without becoming so random that emergent patterns could not be perceived or interpreted. A distinction can be made between "semantically unspecified" and "semantically ambiguous" systems. The former promises the most latitude to users in determining their own meanings, but risks failing to afford any meaningful semantic relations. The latter requires more sensitivity in how it indicates openness for interpretation [11].

Figure 2. The History Tablecloth [11] is highlighting the traces of objects in a kitchen.

THE PRESENCE MEDIATOR

The Presence Mediator and its emerging behavior is not meant to convey a single and specific message (the fork moved, so my partner is eating), but rather to create semantic ambiguity that relates to external everyday situations without indicating any judgment about the meaning [20, 11].

The ideal design of the Mediator is based on the given natural cohabiting setting as described in the scenario earlier: It would enable the Mediator to be seamlessly integrated in the environment and share the same appreciation and appropriation of the participant just like any other ordinary domestic object. For example the mediating technology would be invisibly integrated into an object so that it does not influence participants perception and judgements. Pragmatically speaking, that concerns particularly the visual appearance as well as the kind of movement the Mediator entails. Furthermore, each relationship may require a different object all involved participants can relate to. A and B may share something completely different than A and C do for example.

IMPLEMENTATION OF THE PROTOTYPE

Instead of using an existing object and extending it by some alien appearance we decided to go rather for a stand-alone and generic design that allows interpretative appropriation by the participants (Figure 3). It is intended to be appropriated through the ways the participants understand it explicitly as well as through their actions [11]. Moreover, it allows to apply the Mediator across different field tests as described later and all participants to customize it in both appearance and functionality.

Drawing on The History Tablecloth mentioned earlier, The Presence Mediator for this study does not require to be synchronized over distance nor represent the absolute position in both apartments. Within the frame of the research question it is sufficient to simulate this behavior by moving it randomly a few times per day. All the more important is to make the participants understand and believe that both Mediators are representing each other in time and space: Whenever I move my Mediator in my apartment, your Mediator will move exactly the same way at the same time and vice versa.

Figure 3. Stand-alone and generic design for individual appropriation

The final Mediator was made out of medium-density fiberboard (MDF) and cardboard. It was designed in a circular shape and measured 20 cm in diameter and 9 cm in height. The circular shape has a more pragmatic reason in the first place as it supports the movement in case of colliding with adjacent objects. Inside it is equipped with an Arduino and two modified servomotors each equipped with an extended arm. At the bottom of the shell each arm has a slot. If the servo is rotating the arm will hit the ground, lift and move the whole construction. In contrast to wheels arms lift the shell only during the movement and remain otherwise hidden.

Additional sensors could help to capture additional information such as the edge of a table, presence of people or colliding objects. However, as a consequence it would consume disproportionately more power, increase error-proneness and hence rely on the participants attention. By reducing the technical complexity to a reasonable minimum the Mediator was able to last over five days without any maintenance.

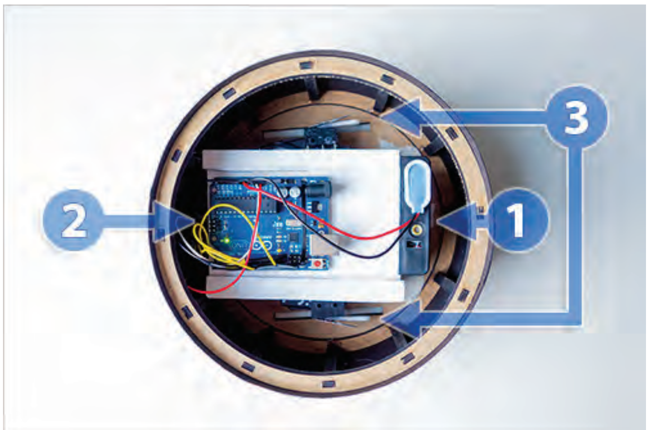

Figure 4. An array of batteries [1] supplies an Arduino micro-controller [2] that controls two servomotors as moving arms [3].

FIELD STUDY

In many areas, where technologies are studied in real-world settings, field research and ethnographical analysis have become established practices in research [11]. The purpose of this study was not only to gain knowledge about the artifact itself, but first and foremost to achieve an understanding of what it means to constitute social presence over time and space, mediated by tangible interaction. In particular this study and its interwoven socio-cultural complexity requires a natural environment.

We conducted two user tests in two apartments in northern Germany each lasting for five days. The first test has taken place between the first author, D, and his apart-living partner (whom we shall call P) in a long-distance relationship. The second test has been performed again in a long-distance relationship between D and his parents-in-law, consisting of the mother (M) and the father (H). All participants were acquainted with common presence technologies such as smartphones, e-mails, social media or voice-over-IP services.

The Mediator was put in the home of the participants. The Mediator was set up in a way that it moved randomly three times per day. Each move varied in direction and distance (0,5m to 1,5m) and took place once in the morning, afternoon and evening. The movement itself was adjusted to a very slow speed and hence noise-reduced.

The participants believed there was a second mediator installed in the home of D. The Mediator was introduced to the participants as minimally as possible. All participants were asked to see the Mediator as an integral part of their chosen everyday object. As far as practically feasible both parts were supposed to be considered as one (mediating) object and not to be detached from each other.

Since this study is about social presence over time and space, no home visits occurred during the study. In both tests all participants had daily obligations and spent most of the day outside the home. It diminished the probability of being home and experience the movement of the Mediator in real-time. That is crucial since the setup of this study is supposed to reflect as much as possible a natural setting of a cohabiting couple, in which a trace of someone else is seen as a trace not when it is created, but only after the fact, as a physical residue of some activity in the past. In contrast, the movement of the Mediator is technologically implemented by servo-motors. Witnessing the thing move could influence participants perception and relation to the Mediator. It would lead to distort the data and most importantly not contributing to the research question.

The role of the participatory investigator

The participatory role of the first author allowed us to conduct both field tests without organizational obstacles and challenges that might have occurred otherwise. The participants and D share a rather private to intimate relationship that is characterized by mutual trust. This position enables D to gain sensitive information and hence insightful data and added to the credibility of the enacted technological set-up. However, being part of the social relation comes with potential subjectivity issues that need to be considered as well. In order to reduce that influencing factor to a minimum all participants were asked to remain silent with D about the study itself during that time period. Furthermore, we decided to apply auto-ethnographical methods in situ to prevent any intrusive and distorting factors. All participants were asked to record the data in any format they feel comfortable with such as video, text or sketches, using an unstructured diary format [5]. Participants were however encouraged to reflect whenever the artifact had been moved and at least once at the end of the day. A final face-to-face interview was added where detailed questions about the diary entires offered a richer and deeper contextual understanding after each field test [16]. In general, however, this study is not meant to provide objective data of peoples behavior, rather than lay bare some relevant design considerations when using tangible interactive objects for social interaction. Subjective as well as objective insights may add to that understanding [18].

In the first field test P (aged 24) lived in a living community together with another even-aged woman. Both shared a kitchen, bathroom and the entrance area with the corridor. If P was at home she has spent most of the time in her own room on around 18 square meters. It was equipped with a bed, a desk, a dresser, a couch and a wardrobe. Furthermore she had a TV as well as a laptop. She decided to use the Mediator as an underlayment for her fruit bowl and placed it on the floor in her room (Figure 5).

The participants of the second field test are a married couple (both mid-50) living in an apartment on about 120 square meters and five rooms. In consultation with W and H it was decided to use the Mediator as an underlayment for their TV guide and to place it on the floor in their living room area (Figure 6). We appreciated the decision as they use the TV guide at least every evening in a social context. In the beginning the wife indicated the main interest in the test and played hence a crucial role.

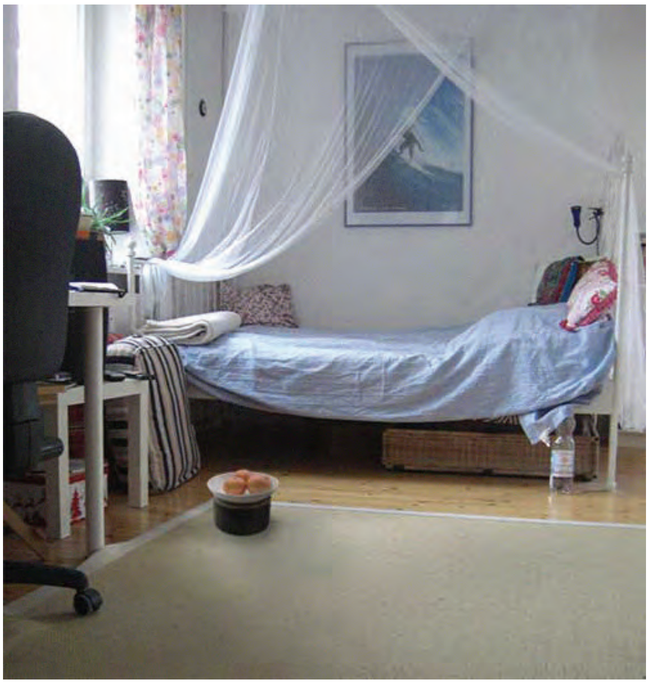

Figure 5. Fruit bowl situated in the participant's room of a living community.

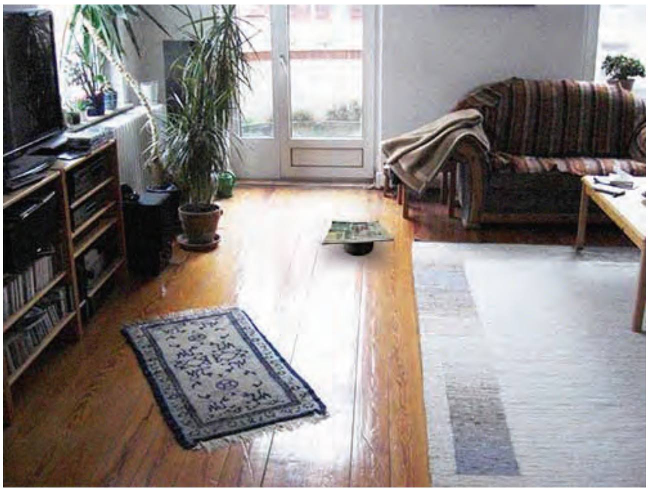

Figure 6. TV guide situated in the living room area of the couple.

RESULTS

First of all, since the setup of this study allows us to do so, we would like to take the issue of D's subjectivity as an opportunity and reflect on his perceptions. Subsequently we will present and discuss the key results in the light of a natural cohabiting setting as to find in a common household. By doing that we elaborate critically and focus particularly on deviant occurrences in that respect.

The role of the researcher as well as a part of the social relation has been a rather unfamiliar situation to D. On the one hand due to the blurring boundaries between private life and university research D is not used to. On the other due to the profession of natural science, namely pharmacy, medicine and biology, all participants pursued. That seems to be reflected in their attitude towards design research practices as D felt initial doubt or skepticism. As it turned out it has been the lack of understanding not only towards this study, but also towards design research in general. Two very different epistemologies, the science tradition and practitioner action, have clashed [1]. All the more important was to explain the very fundamentals and the genuine value of design research to society so that it became graspable and meaningful to them personally. By doing that in a rather informal and leisure setting, unlike the specific instructions at a later stage, D was not just able to counteract the initial confusion but also to diminish the participants potential tendency of feeling morally obliged towards D.

Due to the given restrictions of remaining silent about the Mediator there was hardly exchange about the experiences undergone by the participants throughout the actual study. A few times, however, P leaked some thoughts in terms of gladly confirming a received movement at a particular moment in the past. That feedback put D generally in a positive mood but it came on the same time, due to its intended implementation, with a rather artificial overtone. It has put D morally under pressure. On the one hand because D was not telling the story about the actual implementation and implicitly lying. On the other it demonstrates how fake and artificial concepts can be used to elicit intended emotions [10].

According to our expectations both field tests revealed different insights due to the given context the mediating object was embedded in. While the first field test elicited rather returning entries, the second field test provided a different quality in terms of the discussion that has taken place between the couple. In contrast to the first test, the second focuses more on the co-interpretation of emerging phenomena and provides more breadth in that sense. However, particular results can be extracted from both field tests and hence coincide to some extent.

Mediator versus Detached Agent

Already during the instruction and before the test actually started in both cases the participants assigned their mediating object a name (P as Rudolf; W as Kleiner Freund / Little Friend). That demonstrated the bonding process between the seemingly neutral object and the new owner from the very beginning. Previous research has shown that people tend to treat robots as if they were humans without being actually aware of it [19]. According to [23] perceiving an agent to be human-like has important implications for whether the agent is capable of social influence, accountable for its actions (hence capacity of self-control), and worthy of moral care and consideration. In case of the seemingly abstract appearance of the mediating object, the proclivity for anthropomorphizing can be ascribed to the social environment it is situated in and particularly to the mediating character of the relation D and the participants maintain [22, 23]. It is problematic in a sense that the mediating object, intended to act as a mediator in the relation, obtained its own agency. This relationship is not only between D and the participants. Instead, the mediating object shaped considerably the relation and the experience with its intrinsic properties [6] in a way that it acts as a separate agent in the relationship.

The data reveals this behavior in both field tests. In the first test P ”(...) was thinking about Rudolph and got very excited whether he has moved.” while W was surprised that He covered a huge distance and even mastered the edge of the carpet. I would have liked to see how you managed that little friend. In both cases the participants refer to the mediating object as a detached agent so that the intended mediating properties of the relation were not reflected in their behavior. This behavior is not constant throughout the data set but returning and changing with the intended mediating function. For example, W infers he [the mediator] didn’t move this morning, so D. is probably still sleeping.

Mediator as an Ordinary Object

All participants distinguished to a certain extent, based on the situation, between the Presence Mediator as a means to communicate and the ordinary object as a fruit bowl or TV guide. In the beginning all participants seemed to be very excited and eager about the technology and the seeming fact that it represents each other’s actions. Especially in the first day P had the tendency to deliberately check the fruit bowl for any changes several times. Drawing on the diary entries P stated in the interview that it attracted probably more attention than any other surrounding object or an ordinary fruit bowl. That is ascribed not only to the novelty itself but also to its intrinsic properties [6] that are not given in a comparable ordinary object.

These intrinsic properties got reflected in Ps behavior towards the ordinary object the mediator is attached to as well. P states I have probably eaten more apples than I used to do in the past. By asking P in retrospect, eating more apples may be understood as an implicit excuse to move the fruit bowl and hence the mediating object.

Expectations and Emotions

Throughout the five days the emotional state of P changed almost every time she faced the fruit bowl. Most of the times P was acting between the poles of happiness and disappointment paired with curiosity and expectations: You just made me smile. I was really happy about that [the repositioning] (...) In contrast to positive experiences the noticed absence of movement over a longer period of time provoked rather negative feelings. In connection with additional qualifying attributes P underpinned the emotional intensity: ”I was very very disappointed that it didnt move while I have been at uni.”. This entry explicitly refers to the fruit bowl, however, it seems to be caused by the missing response from D. This demonstrates implicit expectations P adopted over time. Although not explicitly claimed, as a consequence, there evolved a sort of obligation on the opposite side to move the fruit bowl more regularly [24].

Speculation and Creation of Meaning

Especially at the later stage of the test P has been thinking about the fruit bowl outside of the home. P states ”Sometimes on my way back home, I was thinking about Rudolf and got very excited whether he has moved.”

This experience clearly indicates elements of expectation. Paired with curiosity it shapes the basis for pleasant anticipation and provides room for speculation. Once triggered, speculation seems to be an ongoing process and ends once a satisfactory level of meaning is created. Although the fruit bowl relates to external situations, it is still semantically too ambiguous for specific conclusions considering it alone [20, 11]. However, contextual knowledge about the relating partner can play a crucial role in speculation and foster meaningful conclusions. For example, P and D arranged a meeting for the final interview on a Friday at 18:00. P knew the car ride would take D around 1:15 hours. Coincidentally her fruit bowl moved around 16:30 and P concluded that D is still at home. The correlation of gained knowledge is hence food for thought in the creation of coherent meaning.

DISCUSSION

In a way one could say we failed to design what we intended to do: create interactive traces in the physical environment that signal the presence of others over distance. Instead, what we created turned out to be experienced as an explicit object that people attended to and dealt with in a conscious, deliberate mode of interaction. This object sometimes explicitly represented the other person, like an avatar, and at some points we might even conclude it took on an agency of its own, creating a new social relation between the user and the object itself, as illustrated for instance by the fact that the object was given its own proper name.

Nevertheless we gained important insights that can be of relevance to (tangible) interaction design. In particular, this study underscores that it is as of yet an ill-understood challenge to design for what in phenomenology is called "the background" [9]. With this we mean to design an artifact, whose function is to operate in the background of attention, rather than in our explicit awareness [2]. The phenomenological background operates largely unreflectively and immediate, and it is crucial for being able to make sense of what is in our attention at any particular moment. Against this background explicit objects, that are in the center of our attention, stand out [9]. Most product designers, in contrast, would focus on designing such explicit objects, instead of designing their backgrounds. Backgrounds are generally what is not designed, they are "the situation", or "the context", in which the designed object is perceived and acted on. On the other hand, concepts of ubiquitous computing can be said to design for the background, but again in a different way than we mean here. In the ultimate philosophy of ubicomp and ambient intelligence, technology operates completely outside of our awareness, taking care of things without us even knowing it. But when we talk about physical traces of social presence, people do take notice of these traces. They are perceived and acted on, that is, they are taken up within the socio-sensorimotor loop [22]. At the same time, people do not deliberately, and actively, focus on such traces: they are part of the background. In other words, the interactive traces we wish to design for are not completely drawn away from us, but they are certainly also not objects we explicitly attend to. We feel this intermediate kind of objects to be a tremendous challenge for design. Much more work needs to be done to be able to successfully design technology that takes on such a background role. One direction could be to study in more detail how people, instead of physical objects, can themselves at times be part of the background, and at other times figure in our explicit attention, and to see what concrete interactions between people cause these roles to shift.

CONCLUSION

The study demonstrates several discrepancies in contrast to ordinary traces in a natural cohabiting setting. Ordinary objects have been transformed by their added interactive properties. The mediating object shifted from a trace of the activity of the other person, to becoming an autonomous agent with its own intrinsic values. As a result participants dealt with it as an explicit communication tool to maintain exchange in the social relation. This comes with a consequence for the social relation as it entails implicit expectations of exchange the one side and obligations on the other.

In terms of design, the theory of socio-sensorimotor coupling provides a new way of looking at social interaction over distance. At present the theory did not directly provide us with the ultimate answer to how to design for physical traces of social presence (yet). But at least it helped draw attention to the difference between 1) designing explicit representations of other people, 2) designing autonomous agents that take on their own social identity, and 3) designing elements in the overall 'background' by which people make sense of the world. This background function is shown in the way physical traces operate in the ordinary situation to signal social presence. It is as of yet a challenge for interaction designers to achieve the same effect over distance.

REFERENCES

1. Archer, B. The nature of research. In Co-design (1995), 6–13.
2. Bakker, S., Antle, A. N., and Van Den Hoven, E. Embodied metaphors in tangible interaction design. Personal Ubiquitous Comput. 16, 4 (Apr. 2012), 433–449.
3. Clancey, W. J. Situated action: A neuropsychological interpretation. (response to vera and simon). In Cognitive Science, vol. 17 (1993).
4. Clark, A. Being There: Putting Brain, Body, and World Together Again. The MIT Press, 1998.
5. Corti, L. Using diaries in social research. http://sru.soc.surrey.ac.uk/SRU2.html, March 1993. Accessed: 2014-03-10.
6. Dewey, J. Art as Experience. Minton, Balch and Company, 1934.
7. Djajadiningrat, T., Overbeeke, K., and Wensveen, S. But how, donald, tell us how?: On the creation of meaning in interaction design through feedforward and

inherent feedback. In Proceedings of the 4th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques, DIS ’02, ACM (New York, NY, USA, 2002), 285–291.

8. Dourish, P. Where the Action Is: The Foundations of Embodied Interaction, first edition ed. The MIT Press, 2001.
9. Dreyfus, H. Being-in-the-World: A Commentary on Heidegger’s Being and Time, Division I. The MIT Press, 1990.
10. Dunne, A., and Raby, F. The placebo project. In 1. Proceedings of the 4th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques, DIS ’02, ACM (New York, NY, USA, 2002), 9–12.
11. Gaver, W., Bowers, J., Boucher, A., Law, A., Pennington, S., and Villar, N. The history tablecloth: Illuminating domestic activity. In Proceedings of the 6th Conference on Designing Interactive Systems, DIS ’06, ACM (New York, NY, USA, 2006), 199–208.
12. Gibson, J. J. The Ecological Approach to Visual Perception. Houghton Mifflin Harcourt (HMH), 1979.
13. Hornecker, E., and Buur, J. Getting a grip on tangible interaction: A framework on physical space and social interaction. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI ’06, ACM (New York, NY, USA, 2006), 437–446.
14. Hutchins, E. Cognition in the wild. MIT Press, 1995.
15. Ishii, H., and Ullmer, B. Tangible bits: Towards seamless interfaces between people, bits and atoms. In Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems, CHI ’97, ACM (New York, NY, USA, 1997), 234–241.
16. Kenten, C. Narrating oneself: Reflections on the use of solicited diaries with diary interviews. http://www.qualitative-research.net/index.php/fqs/article/view/1314/2 2010. Accessed: 2014-03-10.
17. Kerridge, T., Boucher, A., Law, A., and Gaver, B. Form and movement in domestic networked systems. In Design and semantics of form and movement (DeSForM) (2007), 19–29.
18. Koskinen, I., Zimmerman, J., Binder, T., Redstrom, J., and Wensveen, S. Design Research Through Practice: From the Lab, Field, and Showroom, 1 ed. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2011.
19. Nass, C., and Moon, Y. Machines and mindlessness: Social responses to computers. Journal of Social Issues 56, 1 (2000), 81–103.
20. Sengers, P., and Gaver, B. Staying open to interpretation: Engaging multiple meanings in design and evaluation. In Proceedings of the 6th Conference on Designing Interactive Systems, DIS ’06, ACM (New York, NY, USA, 2006), 99–108.
21. Suchman, L. Plans and Situated Actions: The Problem of Human-Machine Communication (Learning in Doing: Social, Cognitive and Computational Perspectives), 2 ed. Cambridge University Press, 1987.
22. van Dijk, J., van der Lugt, R., and Hummels, C. Beyond distributed representation: Embodied cognition design supporting socio-sensorimotor couplings. In Proceedings of the 8th International Conference on Tangible, Embedded and Embodied Interaction, TEI ’14, ACM (New York, NY, USA, 2013), 181–188.
23. Waytz, A., Epley, N., and Cacioppo, J. T. Social cognition unbound: Insights into anthropomorphism and dehumanization. In Current Directions in Psychological Science, vol. 19 (2010), 58–62.
24. White, M. D. Expectations in relationships: The flip side of obligations. http://www.psychologytoday.com/blog/maybe-its-just-me/201008/ex August 2010. Accessed: 2014-03-10.

Automated Mobile User Experience Measurement: Combining Movement Tracking with App Usage Logging

Lucas P.J.J. Noldus, Ben Loke, Mike Kelia, Andrew J. Spink
Noldus InnovationWorks
Nieuwe Kanaal 5, 6709 PA Wageningen
The Netherlands
l.noldus@noldus.nl

ABSTRACT

We present a new suite of tools for automated measurement of user-system interaction on mobile devices. The system combines measurements of the usage of mobile apps, with a visual representation of the user's location and behavior. This makes it possible to assess not only what the user is doing on the smartphone or tablet, but also what he or she is doing in various different outdoor or indoor locations. The tool offers fully automated data collection and has been designed for deployment in user experience studies with multiple participants for extended periods of time.

Author Keywords

HCI, Android, Mobile, Tracking, User Interaction

ACM Classification Keywords

Tracking, HCI, User Interfaces, Evaluation/methodology, User Interfaces, User-centered design.

INTRODUCTION

The exponential growth of sales of smartphones and tablet computers has been followed by an equally rapid proliferation of software applications ('apps') to run on those devices. However, tools for developers of mobile apps have not evolved at the same pace. One example are tools for measuring the user experience, in particular the interaction of the user with a tablet or smart phone.

User experience and usability of mobile apps can be evaluated in a laboratory setting, by observing a user while operating the device, followed by video annotation and analysis of user-system interaction [1]. However, lab studies of mobile devices are limited in terms of ecological validity, which restricts the scope of conclusions and recommendations. User interaction with mobile devices can also be captured in the field, by recording the interaction using a mobile device camera [2, 3, 4, 5] attached to the tablet of smartphone, with the screen movie (including audio) being stored on a portable hard drive, followed by video annotation and analysis in the lab. Although user interaction can be recorded in great detail, the recording hardware does not permit entirely natural operation of a mobile device, rendering this method unsuitable for observations of long duration. Furthermore, the equipment cost and time needed for data collection prohibit large-scale deployment. An alternative for a mobile device camera is video recording of user-system interaction by a person following the user, as has been proposed for studies in mass transit settings [6] and sport stadiums [7]. However, although this gives the user more freedom to behave naturally, it sacrifices observational detail and is even more labor-intensive than the use of a camera attached to the mobile device. The application of user-centered design methods in mobile app development calls for new tools, offering fully automated user experience evaluation in large numbers of participants.

Several authors have proposed automated data collection methods for remote usability testing on mobile devices. Waterson et al. [8] used proxy-based clickstream logging to capture user traces of a given task without having to modify the participant's software or to access the server. They concluded that this technique allowed them to find usability issues related to the web content displayed on mobile devices. Jensen and Larsen [9] logged user actions on a mobile device in longitudinal studies lasting three months, by instrumenting the app with specific hooks in the source code. This allowed a detailed analysis of the frequency, temporal and sequential structure of the usage of various program functions. However, none of these tools takes into account the user's location during operation of specific apps or functions.

To properly understand how a user makes use of various apps it is necessary not only to obtain data about which apps are in use at which moments, but also, given the mobile nature of the devices, the location of the user needs to be registered. Additional insight can be obtained if we know the movement status of the user: standing still, walking, in a car, bus or train, etc., because this has implications for the ergonomics of the app. If spatial and movement information is available, the interpretation of when, where and how someone is using a particular app has much more power. For example, the function of a public transport app can be assumed to vary according to whether a user is at home (planning a trip) or at a railway station (finding departure time and platform). In order to obtain this information, the designer of a mobile app needs a tool that tracks the location of the user, logs which app(s) are being used, and collects data about the user experience. Examples of questions that may arise are: Where do people use the new

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

shopping planner: at home, on the go or in the supermarket? Does the usage of the app in time and space differ between males and females? Does the duration of usage differ between different locations? Is my app used while the user is walking or only when he/she is stationary?

Noldus InnovationWorks, the research and innovation laboratory of Noldus Information Technology, has embarked on the development of a suite of tools that offer these functions in an automated and integrated manner, to support mobile HCI research and the development of apps for smartphones and tablets. Here we report on the first two components: integrated tracking, visualization and analysis of (a) user location and movement and (b) app usage.

METHODS

The system being developed consists of a number of components: location tracking, movement analysis, app usage logging, data visualization and data analysis. These will be described below. Location tracking and app usage logging are implemented in a Java program named AppTrack™, which runs on smartphones and tablets with Android 2.0 and higher.

useful information. For instance it could enable categorization of the location as 'at home', 'at work', 'commuting' or 'other'. We are also looking into other techniques for indoor positioning, based on video, ultra-wideband, WiFi or ZigBee sensing, to provide more detailed spatial information. The sampling rate (number of fixes per unit time) can be set by the researcher; it is a trade-off between data storage and analysis limitations (higher sampling rate → larger track file), desired spatial accuracy (higher sampling rate → more precise detection of zone entry and exit) and temporal resolution (higher sampling rate → more reliable movement classification), as determined by the research requirements. The default sampling rate is 1 Hz. The track file created on the mobile device consists of a log of fixes, each consisting of date, time, longitude, latitude and elevation in comma-separated values (CSV) format.

The location data collected on the mobile device are passed on, using a custom-design data transfer program using the SMB protocol, to TrackLab™, a PC-based Windows application for track visualization and analysis [9]. This tool was introduced as a research tool in 2013 and is being used for research in psychology, ambient assisted living [10] and consumer behavior [11]. The data are first cleaned up and

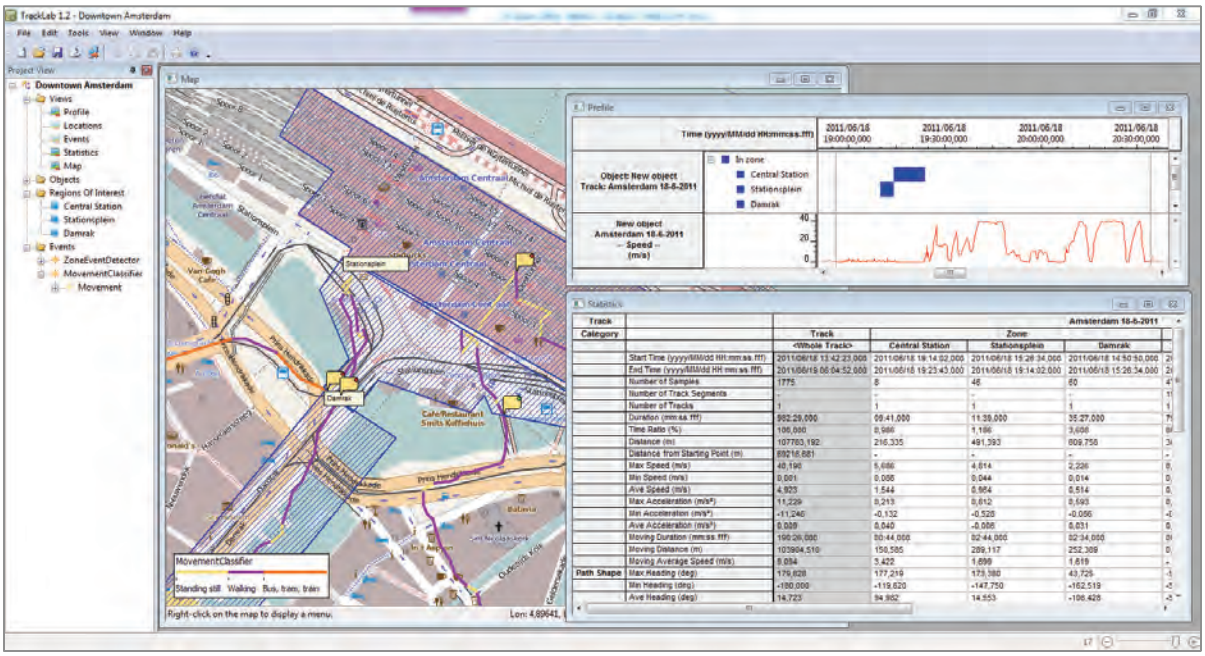

Figure 1. Track of a single user, recorded by the GPS sensor in his mobile device, with color-coded movement categories. The inserts show a profile of zone transitions and movement variations over time, and a table with detailed movement statistics for the entire track and for each region of interest separately.

Location and movement tracking

The built-in sensors of the smartphone or tablet are used to detect and log the spatial position of the user. When the user is outside, or under a roof which does not block GNSS satellite signals, the built-in GPS sensor is used. If a GPS signal is not available, then we will investigate the most practical solution. The rough location data provided by the wireless telecom network will in many cases be sufficiently

filtered using the mechanisms built in to TrackLab, including outlier removal (based on a user-defined acceleration threshold) and smoothing. Especially with GPS data, this is a necessary step due to artifacts inherent in this type of data, especially if the subject is standing still, or if they are in a location with insufficient open sky, or surrounded by reflective surfaces ('urban canyons').

Movement analysis and behavior classification

TrackLab allows movement tracks to be plotted on a map, which can either be an outdoor map, provided by an online map provider, or the floor plan of a building (e.g. an office, museum, hospital, railway station). For online maps, TrackLab currently supports OpenStreetMap but other map providers are considered too. Tracks can either be plotted as individual lines (i.e. one for each participant) or aggregated as ‘heat maps’, which provide a color-graded density plot. Figure 1 shows an example of a track visualization overlaid on a city map.

moment when a particular app gets or loses the focus. At a rate of 0.5 Hz, it generates a list of all active apps, #1 in the list being the app that has the focus. As soon as change in the #1 position occurs, this means that another app has received the focus. At that moment the start time, end time and name of the previous focal app are logged. Going into standby mode (or the screen being switched off) also causes an end time to be logged. The log files are written in CSV format. They are transferred to the PC and into TrackLab in a similar fashion as the track files.

Figure 2. Color-coded app usage overlaid on a movement track. Red = standby mode, light blue = App 1, dark green = App 2.

Based on the position data, TrackLab computes the user’s speed of locomotion. This parameter can be visualized as a color gradient in the track or used to define different categories of movement (e.g. standing still, walking, motorized movement in bus, tram or train). The latter can be visualized as distinct track segments (Figure 1).

TrackLab has been designed to cope with large numbers of tracks, each containing large numbers of data points. The performance is mostly limited by available computing resources. With a sampling rate of 1 fix/second (usually sufficient for spatial event detection and movement classification), a data set can contain tracks of hundreds of participants being tracked for multiple hours.

App usage logging

To record the usage of apps on the mobile devices, AppTrack keeps track of all active apps and records the

The app name in the log file, for one and the same application, can vary depending on the language settings and the smart phone brand. This is a point of attention in studies with multiple participants, if one intends to pool the data across participants.

Data integration

Besides location data, the TrackLab program can also load event data such as the app usage log files, and couple those to the corresponding track file. This opens the way for a detailed analysis of the recorded data from the smartphone, the logging of the active application(s), together with the logged position data. The data can be visualized on an OpenStreetMap map (for outdoor data) or floor plan (for indoor data), such that easily can be seen which apps are used where. The usage of apps in a certain area of interest, a user defined zone, can be analyzed as well. Figure 2 shows a track of a user with color-coded app usage. This visualizes in an intuitive manner how the user alternated between different apps, which app was used where and when, how long each app was used, etc. Detailed statistics can be computed for app usage in user-defined regions of interest (e.g. streets, parks or neighborhoods), time windows (morning, afternoon, evening), or combinations thereof.

DISCUSSION

The rapid development of software for mobile devices requires that tools for measuring the users' interaction with those devices have a number of specific functions. This includes the possibility to measure the user interaction without a keyboard or mouse. Because the devices are mobile, their location and locomotion status of the user are also of critical importance for interpretation of the data. In order to include multiple test participants and collect data for a meaningful period of usage, data collection must be fully automated. The ongoing development presented here is intended as a step in the direction of a comprehensive measurement and analysis tools suite for HCI research on mobile devices and applications.

Future developments

Besides the GPS receiver, smartphones and tablets contain many more sensors that allow the context of use to be recorded, including accelerometers, gyroscopes [12], camera and microphone. Furthermore, we wish to capture the user's intention and experience by providing real-time feedback functions. In the future, these data sources will be added to our tools suite to enhance the capabilities of automated UX evaluation.

ACKNOWLEDGMENTS

The development of the tools presented in this paper is partially funded by the Dutch research program COMMIT.

REFERENCES

1. The Observer XT, www.noldus.com/observer.
2. Weeghmans, L. (2006). User centered design in multimedia broadcasting. *Proc. 15th IST Mobile & Wireless Communications Summit*. Mykonos, Greece: ICCS-NTUA.
3. Wesson, J. (2007). Teaching HCI from an interative design perspective. *Inventivity: Teaching theory, design and innovation in HCI, Proc. HCIEd2006-1 First Joint BCS/IFIP WG13. 1/ICS/EU CONVIVIO HCI Educators' Workshop (Limerick, Ireland, 23-24 March 2006)*.
4. Schusteritsch, R.; Wei, C.Y.; LaRosa, M. (2007). Towards the perfect infrastructure for usability testing on mobile devices. *CHI '07 Extended Abstracts*, 1839-1844.
5. http://www.noldus.com/human-behavior-research/accessories/mobile-device-camera-mdc.
6. Duh, H.B.L.; Tan, G.C.B.; Chen, V.H. (2006). Usability Evaluation for Mobile Device: A comparison of laboratory and field tests. *MobileHCI'06 (Helsinki, Finland, 12-15 Sep 2006)*, 181-186.
7. Sun, X.; May, A. (2013). A comparison of field-based and lab-based experiments to evaluate user experience of personalised mobile devices. *Advances in Human-Computer Interaction*, vol. 2013, article ID 619767.
8. Waterson, S.; Landay, J.A.; Matthews, T. (2002). In the lab and out in the wild: remote web usability testing for mobile devices. *CHI 2002 (Minneapolis, USA, 20-25 April 2002)*, 796-797.
9. http://www.noldus.com/innovationworks/products/tracklab.
10. Noldus, L.P.J.J.; Loke, B.; Spink, A.J.; Bijwaard, D.J.A. Bijwaard (2013). TrackLab: a new tool for indoor and outdoor tracking, movement analysis and behavioral event detection. *Bringing Together Indoor and Outdoor Mobility Solutions: Methodological Approaches for Navigation, Localization, and Environment Modelling. Workshop at AAL Forum 2013 (Norrköping, Sweden, 23 Sep 2013)*.
11. Spink, A.J.; Loke, B.; van der Aa, N.; Noldus, L.P.J.J. (2013). TrackLab – An innovative system for location sensing, customer flow analysis and persuasive information presentation. *PeTRE: Workshop on Pervasive Technologies in Retail Environment, UbiComp'13 (Zurich, Switzerland, 8-12 Sep 2013)*, 985-990.
12. Shoaib, M.; Scholten, J.; Havinga, P.J.M. (2013). Towards physical activity recognition using smartphone sensors. *10th IEEE International Conference on Ubiquitous Intelligence and Computing, UIC 2013 (Vietri sul Mare, Italy, 18-20 Dec 2013)*, 80-87.

WHERE WE LIVE

chi
sparks
2014

Family in Focus: On Design and Field Trial of the Dynamic Collage [DC]

Jasper Jeurens
HAN University of Applied Sciences
Department of Information Technology, Communication and Media
Ruitenberglaan 26, 6826 CC Arnhem, The Netherlands
jasper.jeurens@han.nl

Koen van Turnhout
HAN University of Applied Sciences
Department of Information Technology, Communication and Media
Ruitenberglaan 26, 6826 CC Arnhem, The Netherlands
koen.vanturnhout@han.nl

René Bakker
HAN University of Applied Sciences
Department of Information Technology, Communication and Media
Ruitenberglaan 26, 6826 CC Arnhem, The Netherlands rene.bakker@han.nl

ABSTRACT

In this paper we present the design and field trial of the Dynamic Collage. The Dynamic Collage was designed to facilitate and to stimulate participation of family members in the informal care of an elderly person. The Dynamic Collage enabled relatives to update their current activity by sending a photo to a digital collage at the elderly person's living space. The service is triggered when a family member visits the elderly person. The field trial revealed that all family members valued this type of communication and that they became more aware of informal care. This shows there are opportunities to support informal care in a broader circle than current practices allow. Apart from informal care, our design case contributes to the field of social awareness systems, which we will discuss in the paper.

Author Keywords

Awareness Systems, Persuasive Technology, Healthcare Participation, Informal Care, Photo Sharing

ACM Classification Keywords

H.5.2 [User Interfaces]: *User-centered design*. H.5.3 [Group and Organization Interfaces]: *Computer-supported cooperative work.*

INTRODUCTION

This paper presents work which was executed in the context of current changes in the Dutch healthcare system. Like in other western countries, Dutch healthcare faces challenges related to demographic changes in society - in particular an ageing population - which raises concerns about the sustainability of healthcare in the future. In response, three radical changes in healthcare policy are proposed: the focus of health care should shift from care to prevention, the orchestration of care processes should shift from caregivers to patients, and informal care, such as provided by families, should be supported. From the start, the focus of our work has been informal care, although our results relate to all three changes in health care.

Informal care provided by family members has been the subject of attention for some time. It is argued that it can improve the quality of life for the care dependant elderly, as well as the quality of professional care they receive [1,11,27]. However, some healthcare institutes have difficulty finding concrete ways to stimulate or support informal care; their current work-processes are not organised to this end and the families surrounding a client remain largely unknown.

We explored new avenues for supporting informal care with an opportunity-oriented, user-centered design-research project. This resulted in the design and evaluation of the Dynamic Collage (DC), a photo sharing device for the elderly and their families. The Dynamic Collage addressed the desire of the healthcare institutes present in our quality review boards to increase family participation, and it fitted within well-known HCI design-research traditions such as social awareness systems [13] and persuasive technology [6].

Although the work in these HCI fields was not the starting point of this project and the relevance of these fields to our work - and vice versa – only became clear during the design process, we start this paper with a discussion of this related work. We will focus on the field of awareness systems, returning to persuasion at the very end of the paper. Following the discussion of awareness systems, we present the design case. We give an overview of the design process, the key insights that emerged along the way, and provide a description of the designed system. Next, we discuss the field trial with two test groups. Our findings in this trial bear insights which are informative for the design of awareness systems, as well as for the design of solutions for informal care. We will discuss these findings in the conclusion section of this paper, in which we will also address our future work towards persuasive interfaces.

RELATED WORK

Social Awareness systems

The HCI community became interested in Social Awareness (SA) since the early 1990s [17]. Awareness systems are a class of Computer Mediated Communication systems that

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

support individuals to maintain a peripheral awareness of each other's activity, with low effort and over moderate to longer periods of time [14]. In recent years, many SA systems have been developed. Seminal work for awareness in the workplace was done at Xerox Parc, where Media Spaces were constructed involving screens with video feeds to provide awareness of distant locations and colleagues in an office complex. Since then, researchers have explored other approaches for workplace awareness, and the idea of supporting users with background information about one another [17].

Social Awareness systems focusing on the home and family have shown that such systems can be emotionally supporting, in particular in regards to reassurance and connectedness. The Digital Family Portrait (DFP) [18] for example, focused on reassurance. It recorded the daily activity of an elderly person using sensors in the house. This information was remotely displayed on a portrait picture of the elderly person, with scaling butterflies. Subjects, in a long term evaluation of six weeks, claimed the DFP brought *peace of mind* to the children in the family around the elderly person. *Connectedness* is another goal which received research attention. In particular Astra [14] and Snowglobe [26] showed that awareness systems can improve the connectedness within a family. A trial of WhereAboutsClock [20], where the (coarse-grained) location of family members was displayed on a device resembling a normal clock, showed both effects. The system blended in in the lives of its users, supporting coordination and awareness, with reassurance and improved connectedness as collateral benefits [19].

Much work in awareness systems focuses on automatic capturing of information about its users and sharing this with others. But another path of inquiry explored lightweight communication instead, allowing users to share information intentionally and voluntarily. Astra [14] supported easy within-family communication by photo sharing. Photo sharing was also explored by Biemans and van Dijk [3] and Ashkanasy et. al [2]. Photo sharing has become so widespread with the advent of camera-phones that photo conventions are changing. Photos are used in a much more informal way, almost as a regular form of speech [8]. This brought Biemans and van Dijk [3] to explore functional photo sharing.

A last line of work which is relevant for us, are recent attempts to connect SA systems to existing social media infrastructures. Work on Facebook Listener [25] followed the idea that specific social and business practices could be integrated with social media through *integration software* [24], showing that by linking an SA system with a social network site, the potential range of an awareness system can be increased. Such a coupling of social awareness systems and social media infrastructure may be an antidote to the network effect which generally slows down the adaptation of groupware [7].

Although there is a large amount of work on awareness systems, the field is being criticized for being just another form of technology push. This may be misguided [12], but much of the work we discussed in this section does have the form of 'an awareness system' in search of 'a user need'. Our study follows the opposite direction. It departs from user needs, and shows how an awareness system is a viable solution to address these needs.

DESIGN CASE

1:10:100; three independent, concurrent design cycles

This project was executed in a consortium consisting of our university and three health care institutions. As the original question was quite unarticulated, we have chosen to use the 1:10:100 approach [21, 23]. 1:10:100 is an opportunity-oriented design approach: it supports both the designer as the client to consider new ways of solving the problem [21].

The basic idea behind 1:10:100 is to complete the project in three independent, concurrent cycles with an increasing time span. Each cycle consisted of user-centered design steps needed to complete the project: research, requirements elicitation, creating design solutions and testing the design solutions. We used the Development Oriented Triangulation Framework to ensure balance of research and other efforts within the iterations [22].

After each iteration, user insights and the intermediate solutions were presented to the healthcare institutions involved in our project. In these quality review boards, the design was criticized, requirements were elicited and a new joint design focus was set. This allows recognition of early mistakes and flexibility, and it allows both designer and client to reconsider and discuss the problem and solution spaces. It is considered best practice for the designer in a 1:10:100 to come up with early solutions which are provocative [21].

Design steps

In the first cycle (the '1'), we presented provocative design solutions to the clients, which yielded valuable discussion and feedback. We created goodwill for coming up with something different than 'just another website' and the most important stakeholders were identified: clients (the care dependent elderly persons), family of the clients, health institutions and at a larger distance the government and health insurance companies (Figure 1). We identified families as the single most important target group. This shaped the foundations for the second cycle.

In the second cycle (the '10'), we delved deeper into informal care with observations in healthcare institutes and interviews with entrepreneurs in care giving. This helped contextualizing the needs and problems of family caregivers. A key insight was that family caregivers form a heterogeneous group. We made a subdivision between primary family caregivers and secondary family caregivers. Primary family caregivers have a pro-active attitude, often serving as (legal) contact for the healthcare institutes. Secondary family caregivers are more passive and at a distance. This group is typically visible for the patient and direct care givers, but not for the health institution. Both subgroups experience different types of issues in regards to informal care. For example, the primary family is often under emotional stress as a result of the informal care process, and wishes to disclose their effort to other family members to either delegate tasks together or receive some form of recognition for their effort. The secondary family struggles with lack of time and conflicting responsibilities such as providing care for their own children or family.

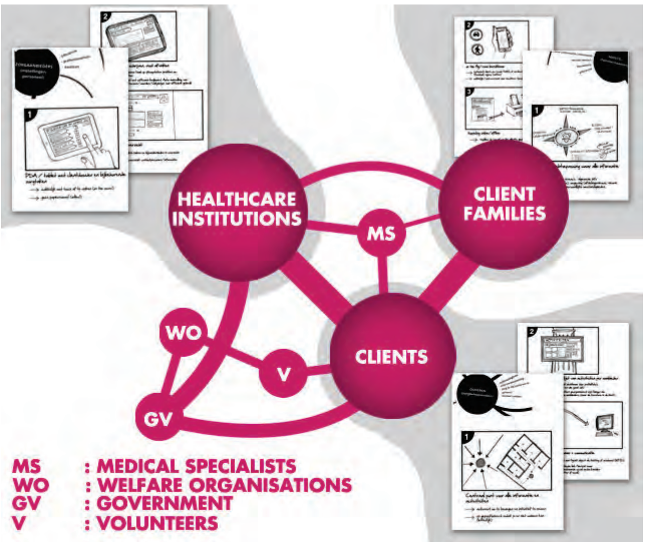

Figure 1: Provocative, low-fidelity scenarios for the three most important user groups in a stakeholder map.

We created a benchmark of existing (online) services for informal care in relation to the needs of the different groups of family caregivers. In effect, existing services try to facilitate (rather than stimulate) informal care processes. This caters specifically to the primary family members who are already carrying most of the informal care load, but excludes the large group of secondary family caregivers.

We completed the second cycle by presenting a design solution that aimed to involve both subgroups (rather than just the 'primary' family members) in the health care process to the clients (Figure 2). The design solution consisted of a 'mood-app' for the secondary group, which supported a more traditional web portal meant to help the primary members of a family. The healthcare institutions recognized the opportunity to include indirect informal family caregivers, and expressed a wish to continue this design direction. A critique on the mood app was that information about mood draws its meaning from a context, which is not visible for the outer subgroup we were aiming for.

In the last cycle (the '100'), participating in the care process as an informal caregiver was further framed. There is a wide array of activities that have an effect on the wellbeing of an elderly in a care institution, which family members can take up. There are practical care giving tasks like cleaning, administering medicine or taking care of laundry for example, but there are also more emotional aspects to care giving. For example, family members can make social calls, take someone out for a walk or trip, or they can simply 'be there' for a person by being physically near. Because most forms of support require someone to be physically near, 'informal care' in light of this paper will be considered as someone visiting the elderly person.

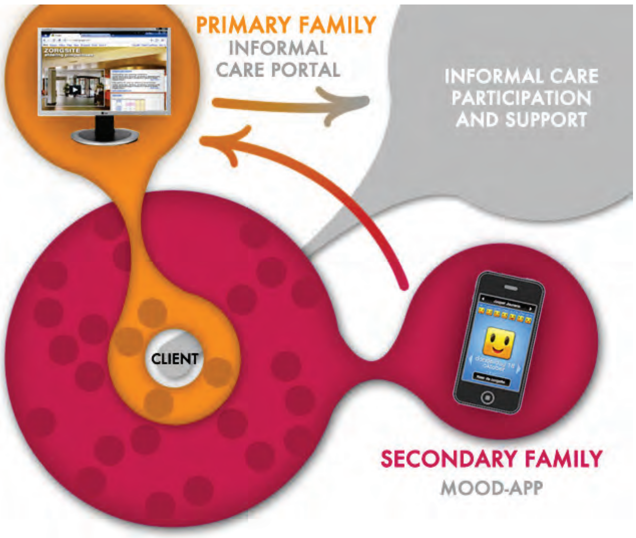

Figure 2: Design of a two-tiered system to include both primary (orange) and secondary (purple) family surrounding a client, through a mood-app and a web portal.

Interviews with family members surrounding an elderly person resulted in the insight that in general, people don't mind helping or taking care of a relative, as long as the relationship is regarded as 'close-tied'.

A framework (Figure 3) was created around the informal care process, identifying four phases that influence decision making in informal care participation. First, the personal background, awareness, emotions and personality traits play a role in care giving: we called this background. Second, family members had to make a decision for an visit or other form of contact. Thirdly, a contact moment took place in one form or the other. Finally, family members made an evaluation or reflection took place.

In a co-creation session, the framework turned out to be a useful tool for discussing the most important motivators in taking part in the targeted behaviour, including existing problems and opportunities for improvement.

The most influential factor on the decision for participation was found to be the nature of the social relationship of the people in the social group. Another effective factor was the presence of a trigger or event.

Final design solution: the DC

Dynamic Collage (DC) is a system where a digital photo frame in the home environment of a care dependent elder person shows a family portrait which is composed of separate pictures sent by family members (Figure 4). The composition of the joined portrait is affected by patterns of participation in the family. Relatives who have participated by visiting will be more visually prominent in the composition. Those pictures are larger and positioned closer to the elder. When members don't show involvement, their part of the composition will suffer in size and opacity (Figure 5).

Figure 3: Framing of the care giving process and its internal and external relations divided in different steps; before (Background and Decision), during (Contact) and after (Reflection).

Family members are allowed to update their part of the portrait only in the event of a trigger, activated by someone visiting the elder person. The duration of the visit is the timeframe in which photo updates are possible. The idea is that such direct feedback encourages social presence. The result will be a composition made of snapshots of 'mundane' moments, all taken at a specific time, rather than the most interesting or precious events in people's lives. After sending the photos, the new composition will be visible on the photo frame at the elderly person, as well as on the mobile phones of the users.

Disclosure of informal care behaviour: creating awareness

By informing all family members when someone is visiting the elderly person, all family members become aware of the informal care behaviour within the family.

We expect that the opportunity to make other family members aware of one's care giving behaviour will positively affect those whose participation behaviour is disproportionally big: the primary family members who put a lot of effort in the informal care process. They benefit from the system because they know their efforts will be noticed within the family. Additionally, by making the informal care behaviour transparent, organizing and discussing the informal care for an elder person becomes easier. We expect that therefore, potential psychological and/or physical problems related to informal care [16] are less likely to remain hidden and can instead be prevented by assisting family members.

Figure 4: the DFC prototype in context of its surroundings.

Figure 5: The composition on the tablet computer. One of the pictures is faded as a result of lacking participation.

Lightweight involvement through contextual photos

Family members have the opportunity to show their involvement (and thereby also acknowledge the informal care behaviour of the person sending the trigger), by sending a photo with their mobile device when they receive a trigger. We expect that secondary family caregivers will want to be involved in this lightweight way. Our interviewees said that they would not mind providing help for a relative but that acts of informal care can easily be discouraged by conflicting responsibilities such as the care for other family members/children, busy work schedules, or practical issues such as the lack of transportation and great physical distance.

We expect that the immediacy of interaction between the person who triggers and the group who sends photos, will contribute to social presence within the family, much as was the case with the WhereAboutsClock [4]. Because of the specific timeframe, we expect the photos to be of a similar nature and context (e.g. pictures in the evening, pictures when the weather is awful) creating a form of social presence.

Supporting behavioural change and outeraction

By giving everyone the responsibility of their own part of the family collage, a sense of joint commitment will give the users a feeling of working together on a common goal with the elderly person as a shared common denominator. These strengthened social ties and group identity could eventually lead to 'outeraction' [15]. Users may want or feel the need to comment or inquire on each other's photos and participation behaviour. Ultimately, the goal is that by keeping the passively natured secondary family members involved and 'in the loop', they will eventually express other ways to contribute to the informal health care process.

TRIAL

Participants

We executed a field trial of the Dynamic Collage with a Wizard of Oz setup. Two test groups were recruited to take part in the trial. We made sure both groups consisted of both primary and secondary informal caregivers in the family. Most participants possessed a camera- or smartphone.

Test group 1 was a family surrounding an 87 year old woman in a nursing home. The eleven participants included all of her four adult children, three of their partners, and all four adult grandchildren. This test group used the Dynamic Collage for six weeks.

Test group 2 was a family surrounding an 88 year old woman who was still living at her own house. At the time of the study the family was debating whether it would be best for her to move to a nursing home, even though she wants to stay at home herself. This family had sixteen participants, including adult children, adult grand children and their partners. One daughter wanted to take part as a primary family member but did not have a Smartphone. Therefore she would only be notified of her informal care behaviour, and would never send photos herself. This test group ran for four weeks.

Setup

We implemented a prototype of the system on tablet computers (a Samsung Galaxy S2 running Android and an iPad 1 running OSX) equipped with SIM to maintain an open connection to the Internet. We used Wakelight on the Samsung tablet to ensure that the automatic sleep mode and darkening of the screen was disabled. The tablets were attached to a folding stand so they could support themselves and be visible from most angles.

A password protected web page hosted the image of the family portrait composition, which was accessed from the tablets by a normal, full screen web browser (Google Chrome and Safari respectively). A script automatically refreshed the webpage hosting the composition when a new picture was uploaded to circumvent the caching features of the browsers. This way we prevented the display of old images of the composition by the web browsers on the tablet. In line with our wish to utilise existing social media infrastructures, we used Facebook groups to relay the composition images back to the senders. To avoid disclosure of personal/family photos, the group was set to secret, so only members of the group could see the content.

Roles

Family members
The participants were asked to provide their telephone number and email address, and asked to notify the researcher by phone when they were visiting the elderly person. Depending on when someone was visiting the elderly person, they would be prompted by a text message. The message contained information about who was visiting and prompted them to use this moment to update their photo on the portrait by sending one to the researcher. The elderly person had no specific task, other than keeping the photo frame plugged in and powered and not using the tablet for other tasks.

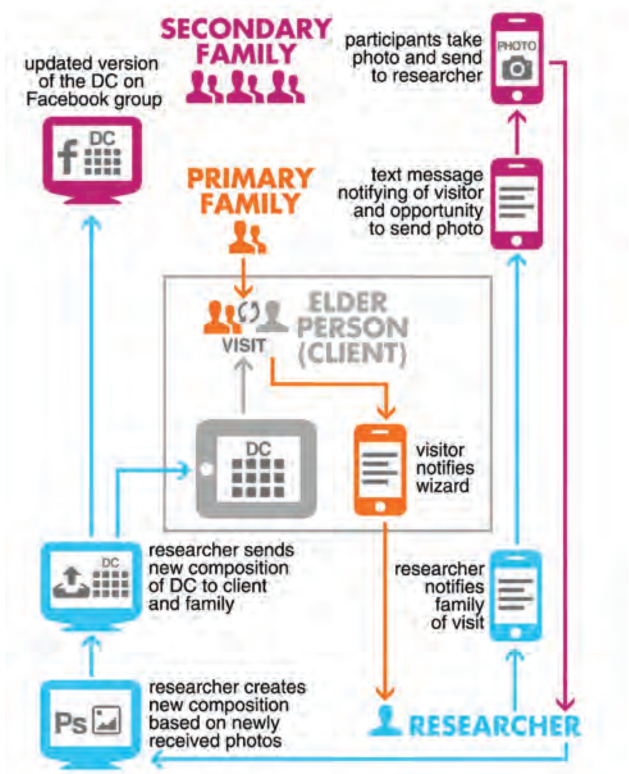

Figure 6: Schematic overview of the field trial, depicting the actions and roles of everyone involved.

Wizard of Oz
One researcher took the role of a Wizard of Oz, relaying the messages around visits to the family behind the scenes. When a family member arrived, the wizard would text a request for a photo to all family members. The Wizard created a single composition in Adobe Photoshop based on everyone's involvement and informal care behaviour, before sending it to the photo frame and the Facebook group.

Data collection

The visits and photo updates were in a spreadsheet. We administered the Affective Benefits and Costs (ABC) Questionnaire [10] to measure changes in social connectedness. Participants filled out the questionnaire in advance as a baseline survey, and again for every two weeks of the trial.

We held weekly semi-structured interviews with different participants of each group. We used these interviews to check if participants had been following instructions, and gather their experiences and opinions about the themes involved in the design. In particular we asked participants about the usability of the system, awareness, disclosure, lightweight involvement, and possible unforeseen effects. The adult children (4 in both test groups) of the elder person were interviewed in a face to face setting in the first week and near the end of the trial. The more distant relatives (grandchildren) were interviewed weekly by phone in the remaining weeks of the trial.

RESULTS

Roles

While test group 1 went without notable problems, test group 2 experienced some start-up problems. Although the care dependent elder of group 2 agreed to participating in the trial (on the premise that no specific action of her was required), she had trouble understanding the system, and kept asking how the tablet would record anything in the household. This troubled her greatly, as she kept asking what would happen if the radio was on, or her neighbour was visiting. This could be the reason the tablet was found unplugged and offline several times by visiting family members in the first two weeks.

With the exception of the second week for test group 2, all visitors reported their visit to the researcher, so the trigger for photo updating could be sent. The trigger was always sent immediately after receiving word of a visit, but in some events, the actual update of the composition back to the frame and the Facebook group was delayed. The Wizard was not always near a laptop or computer (for example, when driving a car or taking a shower). In these cases, the delay was communicated to the participants.

Data

Usage statistics

In both groups, the system was used relatively often in the first week, and less in the following weeks. For group 1 the amount of updates stabilized around 8 updates a week after the first two weeks. For group 2 we see a decline, this trial was too short to see whether usage-patterns would stabilize.

	Group 1		Group 2	
	Visits	Updates	Visits	Updates
Week 1	10	26	4	12
Week 2	5	13	2	12
Week 3	6	8	3	8
Week 4	7	7	4	5
Week 5	4	8		
Week 6	5	9		

Table 1: Amount of updates sent by the groups per week.

Figure 7 shows an overview of family members in terms of visits per week and updates per week. A distinction is made between primary and secondary family caregivers (this is based on who the families claimed to be the primary caregivers, not on an analysis of actual behaviour). It is shown that both primary and secondary caregivers engaged in the trial.

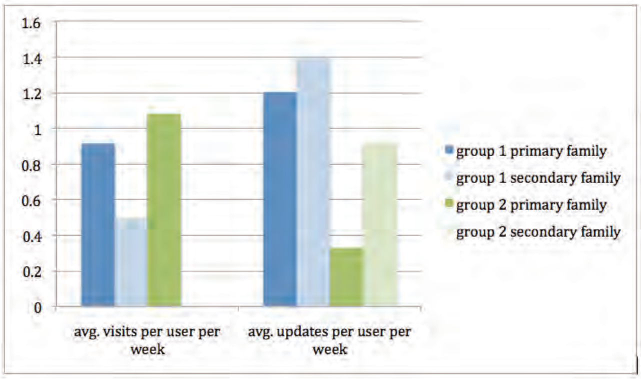

Figure 7: Average activity per user per week. Activity is split between visits and photo updates. Users are split between primary family and secondary family.

In group 2, adaptation of the system among primary and secondary family developed as expected; the primary family, being the children of the care dependant elder person, performed all of the visits, whereas the secondary family participated by sending most of the photo updates. Group 1 showed a similar, though more balanced spread in activity between primary and secondary members.

ABC questionnaires

The ABC questionnaire showed no significant differences over time. We concluded that the questionnaire lacks the sensitivity to reveal changes in such small groups and such small periods of time.

SELECTED INTERVIEW RESULTS

(Group) Awareness

Participants held that their participation in the experiment contributed to awareness of informal care. This was certainly the case for the secondary family members. They made remarks about patterns in the informal care behaviour of others; when asked about their experience of care giving awareness, a grandchild in test group 1 exclaimed "*I suddenly realized holy sh*t, M. is going there 4 times a week".* Such remarks signal the care giving investment of other family members had sunk in.

Also, the care awareness allowed family members to recognize abnormalities in the weekly routine. The partner of one of the daughters in test group 1 explained "*patterns become visible; I could see that it's Monday: Oh, then I guess this and this person will be there again*". These regularities form a necessary background for irregularity, as in the case when one of the primary caregivers in group 1 fell ill; "*I found out about things I would never have known about otherwise, like when M was ill.*"

An unexpected benefit of the system was found in the value of the system for the elderly person and the (primary) caregiver who was visiting. As new family pictures emerged on the Dynamic Collage during the visit, they formed a new topic for discussion. Rather than discussing the goings in the elderly home and the visiting family member, they were stimulated to discuss the broader family and their photos. This was valued positively by the participants.

Lightweight photo sharing

Users showed particular effort to make their photos meaningful. This is in line with previous work by Romero et al. [14] that shows how much people value personally targeted effort and somewhat at odds with the trend that photo sharing is becoming more informal and speech-like [8]. Because of this trend we expected participants, in particular the younger ones, to share fairly mundane updates of 'everyday moments'. However, our participants mentioned that, instead, photos were set up purposely to provide "food for talk' for the visitor and the elderly person. This also meant that participants valued photo updates as a care giving contribution. Several family members from both test groups who were visiting, reported the functional use of photo sharing as helpful and valuable.

However, this value did come at a cost. Photo updates were less lightweight than we intended; participants wanted to make a meaningful contribution (instead of sending a 'random' photo), so the act of updating one's picture became a task of some effort. One of the more active users of the digital portrait reported: "*I did find it increasingly hard to think of something fun to send. I didn't want to send yet another dinner photo when I saw P's*".

The efforts that went into creating a valuable photo for the elderly person fired back to the senders as well; the short turnaround time of the photos in the collage (their presence could be as short as a single day) was a source of some annoyance. Senders felt this didn't reflect the time and effort spent in making the photo: "*I think it's somewhat of a waste to put a lot of effort into [making] a picture when it could be replaced the very next day*".

Social connectedness, group attraction

Photo sharing led to an increased sense of connectedness in the participating families. One of the grandchildren in group 1 mentioned quite literally that the regular sending of photos meant that "*the family was now more connected*".

The forced, specific timeframe to submit photo updates gave all sent photos a 'timestamp' as common attribute. For example, when someone was visiting eight o'clock in the evening, all sent photos would be of an evening atmosphere. This awareness of being involved in the same activities contributed to a sense of unity; "*It was really nice to see all the pictures coming in when everyone was having dinner, it was like we're having dinner together*".

Outeraction

The intended behavioural change of secondary, passive family members to express ways of interacting and involvement outside the system was not notably observed through the use of the system. Although behavioural change could arguably need more time to develop, users reported that the threshold to engage in different ways of communication was simply too high. One direction of future work could be to focus on lowering this threshold, possibly by allowing more interaction within the system.

CONCLUSIONS & DISCUSSION

What started as a project to explore improvement of communication between healthcare institutions, the elderly, and their close families, ended in the design and trial of the Dynamic Collage; a system that aims to support the primary family members involved in informal care, but also to include the broader (secondary) circle of 'distant' family members in informal care through lightweight involvement and (soft) persuasion.

The Dynamic Collage addresses the identified user needs of both groups, including transparency, validation and lightweight effort. Also, the system aimed to increase social connectedness as an important contributor for behavioural change.

A number of conclusions can be drawn from both the design and field test of this system.

Design for secondary informal care

Our study has identified an opportunity for supporting informal care. Rather than focusing on the primary informal caregivers we can aim at solutions which also include secondary informal caregivers. This group is willing to participate in the care giving process but wants to do so in a lightweight manner and this group is in need of proximate incentives that trigger action.

A disadvantage of supporting lightweight participation could be that lightweight contributions (such as photo sharing) is considered less valuable as more meaningful contributions (such as visiting) and that the light contributions might replace meaningful ones. However, we found no evidence for a replacement effect and we found that photo sharing in the way we implemented it, was considered a useful and valuable contribution in itself.

When designing for families or groups, different requirements for different users exist, each requiring their own way interaction with the system. Fisher calls this 'a rich participation ecology' [5]. We have been successful in addressing address the requirements for different users in a single awareness system, uniting different users towards a common goal. Secondary informal caregivers become aware of the activities of their peers and are offered a way to provide support; primary informal caregivers and the elderly become aware of the support of informal caregivers, and the elderly home/institution has an entry to talk about the secondary informal caregiver.

Awareness systems as a design template for informal care

This study suggests that awareness systems can act as the core of a solution for this group of secondary informal caregivers. The data from our design and trial suggests that the Dynamic Collage supports awareness of care giving behaviour in the family, and that users value this increased awareness. As such we consider this a successful design of an awareness system; evidence that there is something to the idea.

But we also faced limitations of the concept. The idea of awareness is information centric: it focuses our thinking on what the information is that people need awareness about [19]. But in our experience, identifying the right information is not a fruitful starting point for the design for awareness. Users may benefit from, or even appreciate certain awareness information and still not seek this information. Users of the Dynamic Collage seem to primarily want to communicate with family members; awareness emerges as a result of this communication. Once people get information, they want to act. Hence, the design of awareness systems quickly dissolves into the broader field of social interaction design.

Photosharing for awareness

Photo sharing turned out to be a flexible, lightweight and fun way to achieve increased awareness and to address latent communication needs. This has also been reported by [2,3,14]. However, in contrast to some emergent practices [8], our participants typically did not want to share photos that were too mundane. Rather they considered taking a photo as a personally targeted effort [14] that should carry its own narrative.

Surprising to us, was the way photos were used as a valuable, remote support for the visitors of the elderly. Our participants valued the possibility to shift the conversations with the elderly person from the goings in the home to the activities of the extended families, those sending photos. This led to a broadening of the narratives with the elderly person, and it may contribute to family cohesion.

Designing for long-term use

We have tested the Dynamic Collage for 4 to 6 weeks. We felt this period was necessary to overcome a novelty effect. Our data suggest we were past the largest novelty effect after two weeks, but it is hard to say how usage patterns would evolve outside of the context of a trial (where continual interviews reminded users to keep on using). There were some signs that design for long-term use demands a system with a richer functionality.

For example users commented on the limited possibilities of a system focusing on photo sharing within a certain time frame. As users considered sending a photo as a personally targeted effort, they did not want to send a similar photo twice. However, within a few weeks this was bound to happen. A more diverse set of exercises or extra incentives could address this need. Another approach is to design a system which is less dependent about the here and now of photo sharing, for example by proving a history function on the tablet.

Design for Persuasion, Future work

We mostly considered the system we presented in this paper as an Awareness System. However, it could also be called a Persuasive Awareness System. The ultimate goal of our system is to make informal care more of a joint exercise among all family members, also outside the direct scope of the system itself.

It seems plausible that increased awareness brings about changes in how families organise informal care, but within our study we have not yet found any evidence that this is indeed the case. It might be that our evaluation lacked the sensitivity to observe these changes, or that the period of 4-6 weeks was too short. However, it could also be that an increased awareness is simply not sufficient for persuasion beyond the action possibilities of the system. We are planning on studies probing into these questions.

ACKNOWLEDGEMENTS

We thank the participants of our trials and the health institutes involved in the 1:10:100 consortium.

REFERENCES

1. ActiZ organisatie voor zorgondernemers, Expertizecentrum informele zorg. (2007). *Samenspel met Mantelzorg* publicatienummer (ActiZ publication no. 07.001). Retrieved from website healthcare expertise center: http://bit.ly/19Z48J5
2. Ashkanasy, S., Benda, P. and Vetere, F. Happy coincidences in designing for social connectedness and play through opportunistic image capture. *In Proceedings of the 2007 conference on Designing for User Experiences*. ACM, 2007, 4.
3. Biemans, M., van Dijk, B., Dadlani, P. and van Halteren, A. Let's stay in touch: sharing photos for restoring social connectedness between rehabilitants, friends and family. *In Proceedings of the 11th international ACM SIGACCESS conference on Computers and accessibility*. ACM, 2009, 179-186.
4. Brown, B., Taylor, A. S., Izadi, S., Sellen, A., Kaye, J. and Eardley, R. Locating family values: A field trial of the Whereabouts Clock. In: *UbiComp 2007: Ubiquitous Computing.* Springer, 2007, 354-371.
5. Fischer, G. Understanding, fostering, and supporting cultures of participation. In: *Interactions,* 18, 3 (2011), 42-53.
6. Fogg, B. J. *Persuasive technology: using computers to change what we think and do.* Ubiquity (2002), 5.
7. Grudin, J. Groupware and social dynamics: eight challenges for developers. In: *Communications of the ACM*, ACM, 37, 1 (1994), 92-105.
8. Hand, M. *Ubiquitous Photography.* Polity, 2012.
9. Hummels, C. and Frens, J. Designing for the unknown: A design process for the future generation of highly interactive systems and products. In: *Proceedings of EPDE.* (2008). , 2008, 204-209.
10. IJsselsteijn, W., van Baren, J., Markopoulos, P., Romero, N. and De Ruyter, B. Measuring affective benefits and costs of mediated awareness: Development and validation of the ABC-questionnaire. In: *Awareness Systems*. Springer, 2009, 473-488.
11. Jones, P. H. *Design for Care. Innovating Healthcare Experience.* Rosenfield, NewYork, USA, 2013.
12. Markopoulos, P. A design framework for awareness systems. In: *Awareness Systems*. Springer, 2009, 49-72.
13. Markopoulos, P., De Ruyter, B. and Mackay, W. *Awareness systems: Advances in theory, methodology, and design.* Springer, 2009.

14. Markopoulos, P., Romero, N., van Baren, J., IJsselsteijn, W., de Ruyter, B. and Farshchian, B. Keeping in touch with the family: home and away with the ASTRA awareness system. In: *CHI'04 extended abstracts on Human factors in computing systems*. ACM, 2004, 1351-1354.
15. Nardi, B. A., Whittaker, S. and Bradner, E. Interaction and outeraction: instant messaging in action. *In Proceedings of the 2000 ACM conference on Computer supported cooperative work.* ACM, 2000, 79-88.
16. Oudijk, D., de Boer, A., Woittiez, I., Timmermans, J., de Klerk, M. *Mantelzorg uit de doeken*. (Sociaal en Cultureel Planbureau), ISBN 9789037704860), 2010. http://www.scp.nl
17. Rittenbruch, M. and McEwan, G. An historical reflection of awareness in collaboration. In: *Awareness Systems*. Springer, 2009, 49-72.
18. Rowan, J. and Mynatt, E. D. Digital family portrait field trial: Support for aging in place. In: *Proceedings of the SIGCHI conference on Human factors in computing systems.* ACM, 2005, 521-530.
19. Schmidt, K. The problem with 'awareness': Introductory remarks on awareness. In: *Computer Supported Cooperative Work (CSCW)*, 11, 3-4 (2002), 285-298.
20. Sellen, A., Taylor, A. S., Brown, B. and Izadi, S. *Supporting family awareness with the whereabouts Clock.* In: Awareness systems: Advances in theory, methodology, and design. Springer, 2009, 425-445.
21. van Turnhout, K., Hoppenbrouwers, S., Jacobs, P., Jeurens, J., Smeenk, W. and Bakker, R. Requirements from the Void: Experiences with 1:10:100. In: *Proceedings of CreaRE 2013, Essen, Germany; 2013.*
22. van Turnhout, K., Craenmehr, S., Holwerda, R., Menijn, M., Zwart, J.P. and Bakker, R. Tradeoffs in Design Research: Development Oriented Triangulation. In: *Proceedings of the 27th International BCS Human Computer Interaction Conference* (HCI 2013), 2013
23. Van Turnhout, K., Leer, S., Ruis, E., Zaad, L., Bakker, R. *UX in the Wild: on Experience Blend & Embedded Media Design.* Presented on The Web & Beyond Conference, 2012, Amsterdam, The Netherlands. (http://bit.ly/SXXYg0)
24. van Turnhout, K., Holwerda, R., Schuszler, P., Tijsma, L. and Bakker, R. Interfacing Between Social Media, Business Processes and Users. A Design Case Study. In: *Online proceedings of Chi Sparks*, 2011.
25. Veen, S., Visser, T. and Keyson, D. V. "I heard you were on Facebook": linking awareness systems to online social networking. In: *Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design.* ACM, 2012, 703-706.
26. Visser, T., Vastenburg, M. and Keyson, D. SnowGlobe: the development of a prototype awareness system for longitudinal field studies. In: *Proceedings of the 8th ACM Conference on Designing Interactive Systems.* ACM, 2010, 426-429.
27. Vilans. '*De bewoners en hun familie zijn ons uitgangspunt.*'. (2011). Retrieved from website familyparticipation Vilans: http://bit.ly/1kvdGjg

chi sparks 2014

Creating Better Health Worker - Patient Interaction Using ICT; Design for Applicability and Acceptance

Arthur Bennis
HAN University
Arnhem
arthur.bennis@han.nl

Dick Lenior
HAN University
Arnhem
dick.lenior@han.nl

ABSTRACT

Case studies in three different healthcare settings were carried out to investigate the extent to which a User Centered Design approach (UCD) increases applicability and acceptance of telecare applications. These settings concerned (a) supervisory tasks, (b) domestic support for the disabled, and (c) remote healthcare diagnostics. The qualitative results of all studies were translatable into a direct positive effect on the *Performance Expectancy* and *Effort Expectancy* constructs of the UTAUT model, thereby increasing the acceptance of the application by the user. A relationship between a UCD approach and increased acceptance is shown. Better substantiation of these results is subject of on-going research.

Author Keywords

Healthcare; UCD; UTAUT; Applicability; Acceptance

ACM Classification Keywords

H.5.2 User Interfaces: User-centered design (D.2.2, H.1.2, I.3.6)

INTRODUCTION

Population aging in developed countries is progressing fast. Because of this, more healthcare will be required, which has to be delivered by a smaller, younger labor force [25]. In the near future it will thus become a necessity to be able to provide more care with less health workers. To cope with these prospects and to anticipate on other expected changes in healthcare (governments making cutbacks in expenditure for healthcare, stricter regulations), the paradigm of direct interaction between health worker and patient is not tenable. Healthcare will need to shift towards an interaction model that is more supported by ICT. ICT can offer the required extensive support and enable health workers to work more efficient.

Such an ICT system is referred to as ***telecare***, the use of telecommunication systems for care where the users are separated in space and/or time [4]. Telecare already offers widely accepted solutions for both social care and healthcare at home [18]. Telecare is also deployed in specialized homes where it enables ICT supported communication.

HCI research on telecare at home identifies barriers to the uptake of telecare [4]. One important component of these barriers is the acceptance level of health workers to use telecare systems. This level is possibly influenced by previous bad experiences with ICT systems due to insufficient training or technology with insufficient usability [20]. The design of new forms of interaction has to meet user expectations in such a way that acceptance occurs almost naturally. Such a design is only possible when users participate during the entire development process; research shows that they are actually willing to [4,13].

Because of this identified need for the introduction of suitable ICT in healthcare, and because a number of healthcare providers requested our assistance, we started an empirical research program around telecare innovations to investigate the extent to which a User Centered Design approach (UCD) increases applicability and acceptance of telecare applications [12]. This was executed in an evaluative and elaborative fashion because we wanted to stay close to practice.

To identify the shift in user acceptance, we use the Unified Theory of Acceptance and Use of Technology (UTAUT) model by Venkatesh, Morris, Davis and Davis [19]. UTAUT has shown its applicability in healthcare settings to investigate acceptance in previous studies [9, 24]. We expect faster acceptance of new technology in healthcare when using UCD because achieving the best fit of the technological means on the tasks of the health workers is central when deploying this approach. The relationship between certain project approaches and user acceptance has been investigated before [6, 23], but no research has been found evaluating acceptance after UCD with UTAUT.

In this study we deploy UTAUT to evaluate the effects of UCD on acceptance and applicability. In future work we aim to have UTAUT strengthen the UCD approach from the very start.

Following the literature review, we present the approaches and findings from three specific case studies in various areas of healthcare: (a) supervisory tasks, (b) domestic support for the disabled, and (c) remote healthcare diagnostics.

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

TELECARE IN LITERATURE

Telecare has shown possibilities for not requiring patients to have a face-to-face consultation with a health worker. Positive results are shown in a study conducted in regional Western Australia [10] where rural emergency departments, only staffed by nurses, reduced the need for transfers of patients. Another Australian study [16] substantiates the cost savings claim gained by the prevention of expensive transfers. Clemins, Coon, Peck, Holloway and Min [3] found that telecare proved to be an effective mode for the provision of diabetes care to rural patients in Montana, USA. Gilmour, Campbell, Loane, Esmail, Griffiths, Roland and Wootton [7] showed the potential of telecare for dermatology cases. Based on the results, 50% of the patients could have been managed with a single video conferenced teleconsultation without any requirement for further specialist intervention. In a more recent study, Rezende, Tavares, Alves, Dos Santos and De Melo [15] showed that telecare in Brazil from 2004 till 2010 prevented the physical referral of patients in 64.2% of cases. Successful application of telecare in triage is shown by Wallace [21], be it with the comment that more research is needed. Further literature shows successful application of telecare in areas including teledermatology [22], wound care [8], teledentistry [5], and palliative care [11].

Research has been done regarding current telecare practices. Significant added value was shown, but also a number of concerns became apparent. Barlow, Singh, Bayer and Curry [1] show that, based on the evidence they reviewed, the most effective telecare interventions appear to be automated vital signs monitoring (for reducing health service use) and telephone follow-up by nurses (for improving clinical indicators and reducing health service use). The cost-effectiveness of these interventions was less certain.

Miskelly [14] shows that new technological developments in care at home are likely to make an important contribution to the care of elderly people in institutions and at home. Video-monitoring, remote health monitoring, electronic sensors and equipment such as fall detectors, door monitors, bed alerts, pressure mats and smoke and heat alarms can improve older people's safety, security and ability to cope at home. Miskelly also found that care at home is often preferable to patients and is usually less expensive for care providers than institutional alternatives

Clark and McGee-Lennon [4] state that, despite the current advantages in technology and networking in the home, telecare solutions have not yet been taken up as eagerly as might have been expected. Authors identified existing barriers to the successful uptake of telecare and derived a set of recommendations for the design and implementation of future home care / telecare technologies. These recommendations include findings like deploying a User Centered Design approach and making the technology able to cope with multiple users. A number of ethical issues were identified and included, among others, fear of technology failing and uncertainty about the security of health data. One of the most prevalent findings is that there is a clear demand for awareness raising and knowledge building on the range, scope, capabilities and acceptance of telecare.

Figure 1 illustrates a sample telecare system as presented by Turner and Maternaghan [17]. This model clearly shows the possible extensive network of technologies enclosing the telecare system. The case studies reported here are all executable within this model. We investigate the interaction of the health worker or the patient with the central box in the model – the Telecare system. Is the devised technical support acceptable?

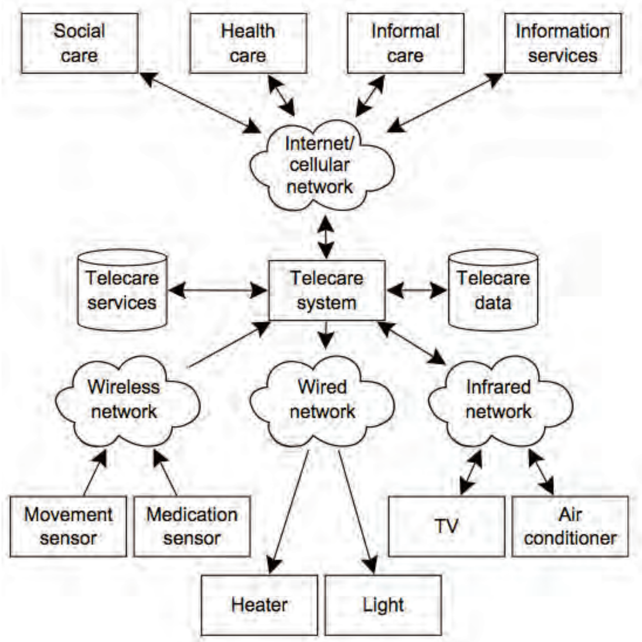

Figure 1. Sample telecare system

To be fully accepted by users, a telecare system has to meet a certain number of expectations of the users. To investigate whether or not the willingness to accept telecare has increased among our subjects, we use constructs of the UTAUT model (see figure 2). These are:

Performance expectancy (PE): The belief of a individual that the system will aid in better job performance.

Effort expectancy (EE): The believed ease of use of a system.

Social influence (SI): The perception of the opinion of important others that the individual should use the system.

Facilitation condition (FC): The belief of an individual that the organization can in fact support use of the system.

Changes in constructs 1, 2, and 3 have a direct effect on Behavioral Intention. Together with construct 4, this Behavioral Intention influences Use Behavior. During these studies we strived for representativeness in the experimental work but did not explicitly take the moderators Gender, Age, Experience, and Voluntariness of Use into account in the field studies.

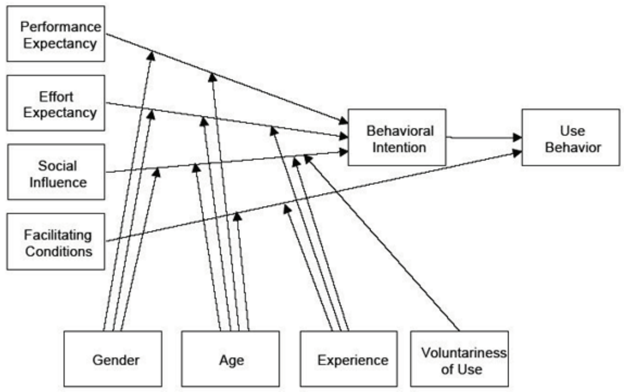

Figure 2. UTAUT model

CASE STUDIES AND RESULTS

In the case studies a UCD approach was deployed to achieve a positive effect on *Performance Expectancy* and *Effort Expectancy* from the UTAUT model, thereby directly influencing *Behavioral Intention*. In each case study the UCD approach was executed in the most conventional way, closely following Maguires interpretation [12]. After a planning phase, an iterative processes containing context of use, requirements, design solutions, and evaluation stages starts. When the solution meets the requirements, the approach is finished. In each case study a careful selection of the available usability methods was made to achieve the best results, therefore not all case studies deployed the same methods in each stage.

Case studies in the field of supervisory tasks

To determine possible technological support during supervisory tasks, and the acceptance of this technology, we investigated remote supervision in psycho-geriatric nursing home wards (living rooms). The health worker legally has to maintain continues supervision over the group, even when admitting occasional care to individuals (guided visits to the bathroom, assistance in acquiring drinks or food e.g.). Through a UCD approach we investigated the best fitting technology in consultation with the health workers.

In a first experiment we investigated sensors to assist the health worker in maintaining supervision. Currently, most deployed sensors in the field of psycho-geriatric supervision are for access control / runaway risk during daytime and for detecting bed leave during nighttime. Health workers often do not experience these sensors to have supporting value in the case of supervision. Our research further shows that these sensors are not well understood (the runaway sensor), they show no direct added value (access control), and they generate too much false positives (bed sensors).

In a laboratory setting, a ward was simulated and six actors where placed inside with a script that was deduced from real world observations. In the first experiment possible combinations of two video camera's, two microphones and two tracking systems (UbiSense and EagleVision) were tested to support in supervisory tasks. In the second experiment combinations of two webcams, a Kinect camera and a depth camera were tested.

From the evaluation with the health workers it appeared that none of the detectors offered a good solution for the supervision task. The most important conclusion deducible from these experiments is that the knowledge and the assessment of the health worker is essential when trying to recognize and interpret incidents. The behavior of patients is very unpredictable and only the health workers that have day-to-day interaction with the patients can assess whether or not a situation is, or will lead to, an incident. Automated detection of incidents by advanced technology is not only very expensive, it is also not effective yet. The interpretation of an incident should for now remain with the health worker.

We did continue to work with audio and video equipment but in a much more focused way, using technology to "extend" the eyes and the ears of the health workers. Audio levels above a certain threshold are clearly a trigger indicating a possible incident. The health worker, busy with tasks outside the ward, detects this noise via a hearing device and is offered a possibility to visually assess the situation via a mobile viewing device. Then the health worker can make a substantiated assessment whether or not to intervene.

To have health workers evaluate the results of such a system, we once again instructed actors to re-enact a number of incidents in a ward. These incident were: (1) patients having verbal and physical arguments, (2) a patient has fallen down; another patient is trying to help him to his feet, (3) a patient has dropped his cup, (4) a patient is calling the health worker because an other patient is pulling him from his chair, and (5) a patient is choking.

All incidents were successfully detected and interpreted using the system; be it with some false positives. The subjects were positive about the outcomes of the experiments and expressed an expected feeling of security and control, even when not physically present in the ward. A feeling of relief of work related stress was expected to be experienced because of the direct available information from the ward.

Case studies in domestic support for the disabled

Are patients able to work with telecare? During our research we identified and solved issues regarding technological support to patients in cooperation with two special patient groups. One of the groups consisted of patients with a mild intellectual disability and the other group was formed out of patients with acquired brain injury. Both investigations showed us that a UCD approach for interaction is indispensable, especially when working with special user groups.

Research was carried out to investigate the requirements for an application to support the acquired brain injury user group with cooking. The cooking process was completely broken down into steps and each individual step was adapted to the capabilities of the users with, among others, progress monitoring and possible alarms. Based on the findings from the various usability tests on the prototypes, development of a production ready application has been started. We observed proud patients with a regained ability to prepare a complete meal.

Deploying a similar approach, it was made possible for users with a mild intellectual disability to independently order their own supper through an application running on a special tablet. Little training on the application was found to be required. A number of open issues were identified which will be resolved to work towards a production ready application.

Both user groups experienced increased personal capabilities when preparing or ordering food by means of the developed technical aids.

Case studies in remote healthcare diagnostics

In the third kind of case studies we investigated the user needs among general practitioners (GPs) regarding a telecare system for remote healthcare diagnostics [26]. From the results of the questionnaires and interviews we can conclude that GPs are initially rather reserved about the possibilities of telecare in their practice. During the interviews, talking about examples, these opinions shifted. After the interviews still eight out of sixteen GPs stated that initial diagnostics over telecare is unrealistic. However, possibilities for follow-up consultations were mostly assessed as feasible. Most GPs feared missing vital information when questioning the patient over telecare and they felt that planning would be an issue. Supporting these reserves in suggesting professional support at the side of the patient during telecare and establishing a virtual waiting room for patients, GPs were less hesitance, but indicated that telecare would only suffice for certain patients with certain complaints. The used telecare system also contains an overview of all the health workers related to a patient; this was judged as very useful by GPs. Also, the possibility to see previous telecare occasions is deemed as very important by eleven out of fourteen GPs.

Next to the research at the GPs, we also worked with a group of health interns on better usability of the telecare application.

Finally, we conducted a small verification experiment with one GP conducting a double consultation of five patients. The experiment confirmed that telecare is usable as a remote healthcare diagnostic tool but especially for uncomplicated follow-up consultations and for intake consultations. The added value of telecare over consultation by telephone is the more personal contact because patient and GP see each other. Patients show reservation to share emotionally sensitive information via telecare.

The GP further believed that a better assessment can be made with improved image quality. In parallel, we investigated, in a more fundamental and experimental study, the effect of mediation on perceived image quality. See [2] for the results.

CONCLUSION AND DISCUSSION

The research reported here consisted of case studies in three kinds of healthcare settings combined with validation experiments in a laboratory. The three settings concerned: (a) supervisory tasks, (b) domestic support for the disabled, and (c) remote healthcare diagnostics. All three case studies showed that a UCD approach has positive impact on applicability and acceptance. All health workers showed increased confidence in task performance.

In (a) the subjects expressed expectations regarding a feeling of security and control, even when not physically present in the ward. Further, a feeling of relief of work related stress was expected to be experienced because of the direct available information from the ward. This shows a positive effect on the UTAUT constructs *Performance Expectancy* and *Effort Expectancy.*

In (b) both the user groups experienced increased personal capabilities because of the developed applications when preparing or ordering food. A positive effect on *Performance Expectancy*, *Effort Expectancy* and *Facilitating Conditions* is shown.

In (c) we observed GPs showing less hesitation to use telecare when examples were discussed or working prototypes were shown. The verification experiment with the GP showed that telecare is usable as a remote healthcare diagnostic tool but especially for uncomplicated follow-up consultations and for some intake consultations. Both these conclusions have a positive effect on *Performance Expectancy* and *Effort Expectancy*.

All our case studies show a connection between at least two of the constructs of UTAUT (*Performance Expectancy* and *Effort Expectancy*) and the organization of the development process using UCD. We did not investigate a possible relation with *Social Influence* and experienced only a minor possible relation with *Facilitating Conditions*. We do believe that these relations also exist and will focus on them in future research. Further, when *Performance Expectancy* and *Effort Expectancy* are affected in a strong, positive way, we expect a resulting positive influence on *Social Influence.*

Each of the three case studies was concluded with qualitative evaluations to assess the effect of UCD on the constructs of UTAUT. Future research should strengthen this demonstrated relation by, for example, deploying validated questionnaires to better measure acceptance effect using UTAUT. Future research should also investigate the degree of acceptance and applicability of telecare applications that are introduced without application of UCD. Are the effects that are observed through UTAUT indeed addressable to UCD?

ACKNOWLEDGMENTS

We would like to thank the student researchers that contributed to this paper with their research: Alex Teulings, Jules Verdijk, Rob Huberts, Rian Engels, and Eva Troost.

REFERENCES

1. Barlow, J., Singh, D., Bayer, S., & Curry, R. (2007). A systematic review of the benefits of home telecare for

frail elderly people and those with long-term conditions. Journal of Telemedicine and Telecare, 13(4), 172-179.

2. Bennis, A., Landman, R. & Lenior, D. (2014). CCTV mediated observation versus non-mediated observation: investigating perceived image quality with different test systems. To be presented at the Human Factors in Organisational Design and Management Conference in 2014
3. Ciemins, E., Coon, P., Peck, R., Holloway, B., & Min, S. J. (2011). Using Telehealth to provide diabetes care to patients in rural Montana: Findings from the Promoting Realistic Individual Self-Management Program. TELEMEDICINE and e-HEALTH, 17(8), 596-602.
4. Clark, J., & McGee-Lennon, M. (2011). A stakeholder-centred exploration of the current barriers to the uptake of home care technology in the UK. *Journal of Assistive Technologies*, *5*(1), 12-25.
5. Daniel, S. J., Wu, L., & Kumar, S. (2013). Teledentistry: A Systematic Review of Clinical Outcomes, Utilization and Costs. American Dental Hygienists Association, 87(6), 345-352.
6. Esser, P. E., & Goossens, R. H. (2009). A framework for the design of user-centred teleconsulting systems. Journal of telemedicine and telecare, 15(1), 32-39.
7. Gilmour, E., Campbell, S. M., Loane, M. A., Esmail, A., Griffiths, C. E. M., Roland, M. O., ... & Wootton, A. (1998). Comparison of teleconsultations and face-to-face consultations: preliminary results of a United Kingdom multicentre teledermatology study. *British Journal of Dermatology*, *139*(1), 81-87.
8. Gray, L. C., Armfield, N. R., & Smith, A. C. (2010). Telemedicine for wound care: Current practice and future potential. Primary Intention The Australian Journal Of Wound Management, 18(4), 158.
9. Hennington, A. H., & Janz, B. D. (2007). Information Systems and healthcare XVI: physician adoption of electronic medical records: applying the UTAUT model in a healthcare context. Communications of the Association for Information Systems, 19(5), 60-80.
10. Herrington, G., Zardins, Y., & Hamilton, A. (2013). A pilot trial of emergency telemedicine in regional Western Australia. Journal of telemedicine and telecare, 19(7), 430-433.
11. Kidd, L., Cayless, S., Johnston, B., & Wengstrom, Y. (2010). Telehealth in palliative care in the UK: a review of the evidence. Journal of Telemedicine and Telecare, 16(7), 394-402.
12. Maguire, M. (2001). Methods to support human-centred design. International journal of human-computer studies, 55(4), 587-634.
13. Martikainen, S., Viitanen, J., Korpela, M., & Lääveri, T. (2012). Physicians' experiences of participation in healthcare IT development in Finland: Willing but not able. International Journal of Medical Informatics, 81(2), 98-113.
14. Miskelly, F. G. (2001). Assistive technology in elderly care. Age and ageing, 30(6), 455-458.
15. Rezende, E. J. C., Tavares, E. C., Alves, H. J., dos Santos, A. D. F., & de Melo, M. D. C. B. (2013). Teleconsultations in Public Primary Care Units of the City of Belo Horizonte, Brazil: Profile of Patients and Physicians. Telemedicine and e-Health, 19(8), 613-618.
16. Thaker, D. A., Monypenny, R., Olver, I., & Sabesan, S. (2013). Cost savings from a telemedicine model of care in northern Queensland, Australia. The Medical Journal of Australia, 199(6), 414-417.
17. Turner, K. J., & Maternaghan, C. (2012). 2. Home Care Systems. Advances in Home Care Technologies: Results of the MATCH Project, 31, 11.
18. Turner, K. J., & McGee-Lennon, M. R. (2013). Advances in telecare over the past 10 years.
19. Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS quarterly, 425-478.
20. Viitanen, J., Hyppönen, H., Lääveri, T., Vänskä, J., Reponen, J., & Winblad, I. (2011). National questionnaire study on clinical ICT systems proofs: Physicians suffer from poor usability. International journal of medical informatics, 80(10), 708-725.
21. Wallace, D. L., Hussain, A., Khan, N., & Wilson, Y. T. (2012). A systematic review of the evidence for telemedicine in burn care: with a UK perspective. Burns, 38(4), 465-480.
22. Warshaw, E. M., Hillman, Y. J., Greer, N. L., Hagel, E. M., MacDonald, R., Rutks, I. R., & Wilt, T. J. (2011). Teledermatology for diagnosis and management of skin conditions: a systematic review. Journal of the American Academy of Dermatology, 64(4), 759-772.
23. Widya, A., Bults, R. G., Sandsjö, L., Hermens, H. J., & Vollenbroek-Hutten, M. M. (2010). A scenario guideline for designing new teletreatments: a multidisciplinary approach. Journal of Telemedicine and Telecare, 16(6), 302-307.
24. Wills, M. J., El-Gayar, O. F., & Bennett, D. (2008). Examining healthcare professionals' acceptance of electronic medical records using UTAUT. Issues in Information Systems, 9(2), 396-401.
25. World Population Ageing. Retrieved March 6, 2014, from http://www.un.org/esa/population/publications/worldageing19502050/
26. Zorgplatform by Zorgservicepuntplus. Retrieved March 6, 2014, from http://www.zorgservicepuntplus.nl/

Experience Centered Design of Energy Interventions for Shared Student Accommodation

Conor Linehan, Derek Foster, Shaun Lawson
University of Lincoln, UK
{clinehan, defoster, slawson}
@lincoln.ac.uk

Maureen Schoonheyt
The Hague University of Applied Science,
The Hague, Netherlands
maureenschoonheyt@hhs.nl

Katrin Heintze
University of Duisburg-Essen,
Duisburg, Germany
katrin.heintze@stud.uni-due.de

ABSTRACT

Delivering effective interventions to motivate people living in shared, pre-paid rented accommodation to reduce their energy consumption is a well-recognized challenge, since there are no financial incentives for people to engage with such efforts. This paper reports on the experience centred design of digital energy interventions for shared student accommodation, led by 100 participant researchers (all students of an undergraduate HCI module), who themselves recruited approximately 300 participants to engage in interviews and design tasks. The research method was informed by principles of participatory design and practitioner-led inquiry, with the intention of eliciting practical, reflective, experiential data to inform the design process. A thematic analysis was carried out to identify clusters of experiences, perceptions, attitudes, behaviours, challenges, and opportunities identified by participant researchers. Findings emphasise the complex social and personal experience of students in interacting with energy consuming devices, and illustrates the value of engaging with these issues at an experiential level.

Author Keywords

Sustainability, persuasive, experience centered design

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI)

INTRODUCTION

The design of technology that facilitates and promotes more environmentally sustainable behaviour has become a major area of focus for the HCI research community [9]. Much recent research explores the possibility of motivating and facilitating end users to make behavioural changes in their use of energy, through the development of digital interventions that allow participants to more easily understand and track their energy consumption. However, there is little evidence of these developments bringing about sustained long-term change in energy consumption behaviour, and little coherence between the conceptual and methodological approaches underlying such work more generally [9]. In order to better inform the design process, this paper eschews further theory building, and instead presents a collage describing the breadth and depth of experiences that participants identify as relevant.

In the following sections, we first present a review of previous work on the design of technology-led energy interventions, focusing specifically on interventions for shared student accommodation. Experience centered design is introduced as a means for understanding the complex social and personal factors that may affect student engagement with energy reductions, while the methods of participatory design and practitioner-based inquiry are discussed as a means for eliciting and understanding participant experience in this context. A thematic analysis of data collected and reflected upon by 100 participant researchers is then presented. We finally discuss the implications of this data for the future design of interactive digital, or indeed non-digital, interventions to motivate more environmentally sustainable behavior in shared pre-paid accommodation.

BACKGROUND

Technology led energy interventions

Research suggests that people typically have a poor understanding of their electricity consumption, since such consumption is invisible, often obfuscated, and difficult to understand [1,4,20]. However, due to the recent proliferation of smart meter infrastructure, it is now possible to develop digital systems that expose energy consumption, and provide interpretive feedback, to users. Thus, the design of technology to help users make behaviour-based reductions in their energy consumption has received a great deal of attention in recent years by the HCI, ubicomp and related communities (i.e., [4,13,5,7]). Furthermore, decades worth of studies carried out by environmental psychologists have demonstrated that providing people with direct, intuitive feedback on their energy use can motivate them to reduce consumption (see [4,9] for reviews). Previous interaction design work has delivered feedback through interfaces such as phone applications [22], ambient displays [12], and social networking sites [7].

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

Energy interventions in shared student accommodation

HCI research into the design of energy feedback systems tends to focus on applications in either the domestic [5] or organizational [8] setting. Different challenges have emerged from these environments; for example, people are unwilling to sacrifice hygiene or comfort related behaviours at home, and report a lack of responsibility for consumption at work. Interestingly, shared student accommodation appears to offer a combination of the challenges of those contexts; student accommodation is a student's home, but due to electricity being charged at a 'flat' pre-paid rate, there are no financial consequences of saving, or indeed wasting, electricity [17]. Indeed, student energy related behaviour has been identified as an issue of significant concern, as evidenced by various campaigns to promote more ecologically responsible behaviour, such as the "student switch off" (http://www.studentswitchoff.org/) campaign in the UK.

A limited amount of previous HCI research has investigated the design of energy interventions in shared student accommodation. For example, Odom et al., [17] evaluated a feedback interface that facilitated competition between separate dormitory buildings, concluding that social incentives proved more motivating to students than environmental concerns. [1] identified variations in energy usage across students, suggesting that targeting the behaviour of a small number of abnormally high consumption students could facilitate significant savings.

The theory-practice gap

Regardless of the context in which they have been implemented, or the design strategy adopted, there are very few examples in the HCI literature of technology-led interventions either sustaining user engagement or, more importantly, facilitating long-term reductions in energy usage [5,9]. Some researchers (i.e., [8,9,13]) have suggested that the failure of these systems are due to a lack of understanding on the part of designers of the complex ways electricity usage fits within and impacts peoples lived experience. In this respect, there appears to be a disconnection between how technology mediated behaviour change works in theory and in practice (see [19] for a parallel discussion of theory practice gap in social science).

The initial stages of design of energy interventions often involve social science-style research on small groups of potential users, in order to build and adapt "theories" of energy usage. Prototype systems are then built to implement those abstracted concepts. Interestingly, the practice of abstracting data gathered in small-scale user studies to form theories that inform a design process has been much criticized recently. For example, Ghassan & Blythe [11] suggest that this approach is symptomatic of researchers not discriminating the "minor science" of design practitioners with the "royal science" of basic researchers. Gaver, et al., [10] criticise the scientific analysis of user data as blunting the connection between designer and user. Similarly, Olivier and Wallace [18] argue that reducing users' experiences to a set of objective data can diminish our understanding, and subsequent valuing, of human heterogeneity. In our work we intend not to build solutions based on abstracted theories, but through contact with the unfiltered experiences and reflections of participants. Indeed, in the context of designing technology that we expect to impact significantly upon people's lives, it seems especially important to engage with participants at an experiential level [20].

Experience-centered design

In recent years, computing technology has developed from something primarily used in organizations to facilitate commerce, to something that impacts upon all aspects of our personal, social and cultural lives [22, 23]. The study of human interaction with computers has moved away from analyses of human cognitive abilities and interface usability, and towards a more holistic understanding of the complex interactions between technology and the human experience [6]. The process of designing technology based on understanding the subjective experiences of users is referred to as experience centered design [23].

Unsurprisingly, there is little agreement on how best to understand the experiences of people as part of the design process. For example, Forlizzi and Ford [6] proposed *subconscious, cognition, narrative,* and *storytelling* as useful analytical dimensions. Wright and McCarthy [22] identified *emotional, sensual, compositional* and *spatio-temporal* as components of experience. Norman [15] breaks experience into *visceral, behavioural* and *reflective*. Despite this lack of consensus, there does seem to be an overall commitment to understanding interaction from a holistic rather than reductive perspective [2].

In addition to the lack of philosophical consensus, there is also no agreed-upon best-practice research method for sampling experience as part of the design process. However, Wright and McCarthy [23] emphasise that the most fundamental requirement is a commitment to dialogue between designers, users and communities. Hence, research methods drawn from the social sciences seem most appropriate. We have identified two research methods that seem useful for understanding participant experience in the context of the current paper, described below.

Participatory design

Participatory design is a process long established in both research and industry. It is an approach that invites the people who will benefit from, or be impacted by, a technology to participate in its design, with the intention of empowering those communities. There is also an assumption that involving participants in the design process should lead to more acceptable, useable and useful technology [21,23]. The practical act of creating solutions to design challenges can also be seen as a means for eliciting more realistic contributions from participants than is possible with interviews and focus groups, which often produce vague, or unrealistic results, or suffer from effects of social acceptability. Further, due to the potentially intrusive nature of energy interventions we feel compelled to adopt a participatory approach to our design work.

Practitioner based inquiry

Practitioner based inquiry is a method not commonly used in HCI research, but one which seems to offer value in sampling and understanding the experience of participants. It is most commonly employed by practitioners such as nurses [19] and educators [14] in investigating and reflecting on their own practice. Since people who generally have minimal training in research methodology carry it out, the method remains controversial and divisive. It is seen by some as entirely unscientific, on a par with spirituality and witchcraft [14], and by others as the necessary process through which subjective experiences can be understood, and theory can be put into practice [19].

There are parallels between the goals of experience centered design and practitioner-based inquiry; both emphasise pragmatism, subjectivity and experience as necessary and valuable (see [19]). It follows that there may be some value in carrying out experience centered design through a manner influenced by practitioner-based inquiry. Specifically, rather than eliciting requirements from participants through interviews, focus groups and design tasks, it may be useful to encourage participants themselves to undertake this research. There are many advantages to this approach. For example, in line with the motivations for participatory design, the people who will be affected by the technology (i.e., students) are empowered to influence its design. However, in the study presented in this paper, participants are not only empowered through being consulted, or collaborated with, but they fundamentally define the very questions that are asked. Thus, in line with practitioner-based inquiry, the narrative generated from research is inherently and entirely composed of the practical lived experience of participants. We feel entirely justified in conceiving of students as ‘practitioners’ in this context, since their own experience, and that of their peers, is exactly the experience they are researching, and that we are interested in. Additionally, the fact that research is planned and carried out by peer accomplices rather than professional researchers reduces the likelihood of sampling invalid socially acceptable pro-environment opinions. In addition, we feel that carrying out research in such close collaboration with students [15] is a uniquely productive approach to teaching and learning HCI at undergraduate level, and one capable of generating publications for tutors.

The disadvantage of the approach is that it does not allow for the generation of a coherent, generalisable theory of energy usage in student accommodation. It therefore suffers from the same criticisms leveled at wider practitioner-based inquiry (see [14]). However, the intention of the study is not to build a scientifically accurate theory, but to inform design through building a collage from the breadth and depth of experiences that participants tell us are relevant in this context. The advantage of the large sample is in the range and breadth of experiences elicited.

METHOD

Participants

One hundred and three people (ten female) were recruited from an undergraduate HCI module to act as participant researchers (hereafter referred to as researchers). Three of those failed to return any usable data, leaving exactly 100 researchers in the final sample. Those researchers themselves each recruited between two and five participants (hereafter referred to as participants), giving an approximate (due to sometimes vague reporting) participant sample of 300. Importantly, all participants and researchers had some personal experience of shared student accommodation.

Procedure

Researchers engaged in experience centered design activities concurrently with their study of the HCI curriculum. They were initially presented with a design challenge, and used experience centered practices to understand and address that challenge. Researchers initially conducted focus groups in order to elicit requirements. They were given the freedom to employ a variety of techniques within those focus groups. The majority (n=53) employed semi-structured interviews, but questionnaires (n=40), card sorting (n=26), participatory design tasks (n=15), diary studies (n=13), cultural probes (n=4), and “cool walls” (n=4) were also used, and some researchers used more than one technique. Each researcher produced a thematic analysis of their focus group data. Finally, researchers produced a paper-prototype based on the findings of their research. Due to both space restrictions and the focus of this paper we do not present the results of the paper-prototyping task.

Data Analysis

Each of the 100 researchers inductively identified three themes in their focus group data, which they commented on and justified with quotes, typically constituting one paragraph of text per theme. The authors of the current paper took the text produced by those 100 researchers and carried out an inductive thematic analysis. Thus, the data presented here represents both the subjective, experiential information from participants, plus the researchers’ reflections on, and interpretations of, that data. It should be stressed that, since the researchers were, themselves, members of the participant group, their own reflections are valid experiential data. Moreover, since researchers had spent time researching and considering the topic, we would assume that their reflections would be better informed and more useful for the purposes of design than that of a naïve participant.

An inductive thematic analysis was carried out on the data obtained from researchers following the method outlined in [3]. Thematic analysis is useful for analysing large quantities of qualitative data, especially in little understood domains where existing theories and models do not exist. Data was first transposed into a spreadsheet and separated out with one sentence per row. Thus, the unit of analysis was at the sentence level. In total, 1,760 of these units were analysed. Open coding was carried out first. Specifically, a researcher read the data closely, attaching a conceptual label (or ‘code’) to each line of data. A total of 87 codes were generated, which were then grouped together based on conceptual similarity, creating 34 learned abstracted categories. Axial coding was then carried out and abstract categories from open coding were amalgamated to create more defined clusters, referred to below as categories.

Theme	N
Student experience	146
Activities	53
Consumption	53
Defining Coolness	40
Energy consumption	232
Living arrangements	15
Listing devices	26
Essential usage	37
Non-essential usage	30
Wasting electricity	80
Environmental impact	20
Wanting to reduce	24
Barriers to saving	291
Lack of understanding	23
Awareness	77
Economic consequences	67
Forgetfulness	12
Laziness	7
Motivation	42
Comfort and convenience	11
Apathy	52
Behavioural Solutions	209
Lifestyle changes	12
Rewards	73
Punishment	17
Competition	77
Cooperation	30
Design suggestions	450
Platform	74
Simplicity	29
Visualisation	112
Functionality	129
Visual Appearance	56
Privacy Concerns	35
Extravagant requirements	15
Describing approach	26
Linking sentences	322
Unintelligible	84
Total	1760

Table 1. Descriptive analysis of number of occurrences of themes identified in the data.

RESULTS AND DISCUSSION

Five distinct categories were identified, based on the type of information that participant researchers were trying to learn through their research. These categories are; *Understanding Student Experience, Understanding Students Experience with Energy, Barriers to Reducing Consumption, Suggested Behavioural Solutions*, and *Design Suggestions*. A descriptive analysis of the relative occurrences of each theme identified is presented in Table 1. The categories are expanded upon below, with unique themes identified within each category. We have attempted, as much as possible, to tell the story in the words of the researchers themselves.

Category 1: Understanding Student Experience

These data (146 mentions) represent researchers attempts at gaining a general understanding of the subjective experience of undergraduate students living in shared accommodation. Due to space requirements, and in order to focus on experiences more specifically related to their electricity consumption, we have kept discussion of this category to a minimum. Participants in our study reported enjoyment of socialising, drinking alcohol, playing computer games and using social media. However, one finding that does seem potentially useful is that researchers consistently described the student lifestyle as inherently social, and suggests that the attitudes of others, and what they consider cool, can influence their own attitudes and behaviour.

Category 2: Understanding Students Experience of Energy

Data classified in this category (232 mentions) represents researchers' attempts at understanding student's attitudes towards, and experiences of, using electricity in their day-to-day lives. Seven distinct themes were identified in the data; *Describing Living Arrangements, Listing Devices, Essential Usage, Non-essential Usage, Wasting Electricity, Environmental Impact,* and *Wanting to Reduce Consumption*. *Describing Living Arrangements*, focused on understanding how distinct patterns of consumption can be seen within a flat; *"With each student spending the majority of time in their rooms…. each room will have electrical appliances/devices turned on, on standby or charging up"* (P32), *"my bedroom because it's got most of the technology in it"* (P43), *"On the opposite, the energy in the kitchen is used by all flatmates"* (P2). Indeed, across all the data gathered, it seems evident that, *"most of the time the students are at home…. some kind of technology is always being used"* (P80). Some researchers chose to focus on listing the types of devices that students used in their accommodation, *"it appears most students use a variety of devices on a regular basis"* (P84), *"The devices in a student house can build up to about 30"* (P89). These ranged from kitchen appliances *"kettle, cooker, oven, microwave, iron and refrigerator"* (P2), *to "TV, games console, laptop"* (P43) to *"Xbox, DVD Player and I-pads"* (P92) and *"smart phones"* (76). Very little of this data is surprising or unique to students.

Some researchers concentrated on distinguishing *Essential Usage* (i.e., usage that participants would be unwilling to reduce) from *Non-Essential Usage* (i.e., usage that could be potentially targeted by intervention). However, there was little consensus across researchers as to which types of usage fell into each of those categories, *"what is seen as essential differs from person to person"* (P53). As expected, researchers found that a large number of students consider energy usage related to cooking and cleaning to be essential, *"they tend to prioritise more essential kitchen and bathroom items required for everyday life over leisure and entertainment"* (P6), *"all of the students will use the kitchen* *ware for example kettle and microwave,"* (P35), *"Heating and Lighting"* (P53), "*one said he found the shower essential whilst the other said he found the shower non-essential"* (P53). Further researchers found that students consider energy consumption related their studies to be essential; *"they needed it to do their work,"* (P13), *"computer and laptops being a necessity for a university student due to the amounts of research and work needed,"* (P32), *"they use energy as part of their education"* (45). Interestingly, the majority of students reported that the use of electricity for entertainment purposes was also essential, *"no motivation would stop him from interacting with the devices he uses for social computing,"* (P4), *"Students said their PC's were so important because "I need it for socialising, video games and movies"* (P56), *"I live a rather busy lifestyle so mainly things like TV, computer, games console, things like that"* (P43). Given that many *"participants were not willing to change their behaviours or hobbies in order to save energy"* (P13), it is difficult to see where savings in energy consumption can be achieved.

A small number of researchers suggested *Non-Essential* uses of electricity; *"energy that was used for entertainment purpose,"* (P4), *"participants find the entertainment items less importance"* (P6). While this finding is interesting, it is in direct contradiction with that discussed above. Some researchers suggested that savings could be made in the usage of devices simultaneously, "*Using a device as background noise is an unnecessary use of energy"* (P32), *"Participant 4 admitted to "listening to music whilst playing on" his "games console and cooking food as well"* (P40). Perhaps this type of simultaneous usage might be the best target for intervention. Surprisingly, one researcher also found that; *they were happy to give up with such devices as: cooker, oven and hairdryer* (P2), *they would use less ironing, keeping light off and using less heating* (P2).

A number of researchers described their participant's behaviour as *Wasting Electricity*. Participants often leave electrical devices turned on regardless of whether they were being used or not. This ranged from lights; *"both students leave lights on unnecessarily"* (P36), *"Two of the students claimed that they could not sleep without the use of a night light"* (P30), to computers, games consoles and televisions, *"evidence in the research of computers being left on"* (P11) *"they left appliance's on such as their laptops and phones chargers, also including leaving their television on standby"* (P15), *"Participant 1 admitted to using his "phone, iMac, games console, lighting, speakers and headset" at one given time"* (P30). Participants also discussed mis-use of charging, *"My worst habit is having my laptop on charge constantly"* (P27), *"Participants stated they charged devices at least daily"* (P7). The observed tendency to leave devices turned on seems to be driven by convenience and comfort, *"they do this because it saves on time"* (P15),*"participants agreed that they thought it was pointless to turn off their devices because they would use it again some time in the short future"* (P52), *"I put my laptop on sleep because I can just open it up and continue what I'm doing instead of having to boot it up"* (P66), *"students cited their hectic schedules as a key influence in their use of technology"* (P43). Beyond convenience, some researchers reported examples of extravagant usage, *"Participant 3 admitted to using her hair straighteners more than once daily, after lectures at University she would "go home, get something to eat and then straighten" her "hair again to make sure that" she "looked nice*" (P30). *"One participant told me that instead of getting his window fixed he just turned his heating on more often*" (P72), *"Participant A mentioning how they enjoyed the warmth of the heating, but also the fresh air – resulting in heating being pumped into the cold winter air"* (P98), *"Have you ever left the window open and the radiator on? "Er yeah it's the best way to do it"* (P11). In summary, since so much energy seems to be wasted in shared student accommodation, there is great potential for reducing energy consumption through targeting these wasteful behaviours. However, a significant challenge exists, as students feel justified in engaging in these wasteful behaviours, as they aid in their comfort and convenience.

Very little of the data gathered made any mention of the *Environmental Impact* of using electricity. There were some instances where researchers reported positive attitudes towards the environment, *"it helps save the planet so that is something I would consider as cool"* (P48), *"One participant was very passionate that reasons to reduce energy are to reduce impact on nature"* (P77). However, the majority of attitudes expressed towards the environment displayed a lack of interest and responsibility, *"students from the focus groups require convenience over things such as impacts over the environment"* (P50), *"...the amount of electricity we use isn't going to contribute anything to like the overall effect of global warming, so that sort of stuff doesn't really bother me"* (P52), *"All cards relating to climate change considered lame on the cool board"* (P57), *"they care more about money than the environment"* (P12).

Some researchers found evidence of participants *Wanting to Reduce Consumption*; *"all of the students want to reduce their energy used"* (P35), *"both participants thought that energy saving is cool"* (P48), *"all the participants think that reducing energy usage is "quite important"* (P60), *"most students are somewhat conscious to saving energy by turning off devices"* (P84). However, this view was not unanimous, and indeed, there were contrary views expressed; *"people that are most interested in being energy aware are fanatical liberals who protest all the time"* (P11), *"My fear is that this wouldn't increase how much energy they are saving as one of the subjects said that even if they knew their usage it wouldn't stop them from being wasteful"* (P101). There are clearly significant challenges to any intervention intended to persuade students to reduce their energy consumption.

Category 3: Barriers to Reducing Consumption

These data (291 mentions) represent researchers' attempts at understanding student's perceptions of the barriers to reducing their energy consumption. Eight distinct themes were identified in the data; *Lack of Understanding, Awareness, Lack of Economic Consequences, Forgetfulness, Laziness, Motivation, Comfort and Convenience, and Apathy.* Reflecting the findings of previous studies [4], a number of researchers found that participants expressed a *Lack of Understanding* regarding their energy usage; *"students may not be too familiar with existing terminology or whether their current energy consumption level is* *particularly high or low"* (P43) *"there is not a defined scale of how much I should and shouldn't be using"* (P81). One participant suggests that, *"if they were better educated in this area, they would feel more conscious on their energy usage"* (P97). It is clear that there is a problem with the students understanding of energy consumption, and that education may play some role in any intervention developed.

A related issue is the awareness of their own behaviour. Researchers found that participants consistently expressed a lack of awareness of their own *Level of Consumption, the Relative Consumption of Devices,* and of *How to Reduce Consumption. "All three participants said they are not aware on how much energy they use"* (P12), *"they did not get any feedback from the company they are with"* (P15), *"There is no meaningful data accessible to students"* (P22) *"there is no way I can find this information out in my house"* (P36). Discussing Relative Consumption; *"they, usually, do not know how much energy they use by each device"* (P2), *"weren't aware of what items used up how much energy"* (P59),*"We always leave the TV on because we forget, so it'd be interesting to know how much energy that's actually using"* (P52). It seems clear that students living in shared accommodation rarely know how much energy they are using, and that this may contribute to overconsumption.

One of the most consistent themes running throughout the data is the impact that a *Lack of Economic Consequences* has on participants' behaviour. *"Students don't take notice of wasted energy if there is no consequence"* (P56), *"I pay a 'flat rate' which doesn't change whether I'm saving energy or not"* (P4), *"there is no financial incentive to save electricity"* (P17), *"As I do not pay for energy bills, I do not care"* (P37). Interestingly, paying a flat rate for electricity, seems to actually encourage irresponsible over-consumption from students; *"the fact that we don't have to pay just makes us like 'meh, we might as well make the most of it'"* (P52), *"as they already paid for it they might as well use it"* (P76). This is a psychologically significant unintended consequence of the traditional pricing policy for shared student accommodation. Undermining this attitude through technological intervention may prove a difficult task. Indeed, students suggestions for interventions typically involved the imposing of economic consequences, *"if there was a possible money benefit from saving energy then it would be seen as more appealing to students"* (P4), *"if I had to pay for the bills then I would try and use less"* (P48), *"they care more about money than the environment"* (P12), *"I save energy to save money, not the planet"* (P45).

A small number of references were made to *Forgetfulness, Laziness*, a general lack of *Motivation, and Comfort and Convenience* as barriers to reducing energy consumption. *"They often had trouble remembering to turn things off"* (P40), *"the light switch, you always forget to switch if off when you're not in the room"* (P52), *"Some people want to save energy, but simply forget to switch off plugs and devices"* (P94). There may be some advantages in targeting this forgetfulness with an intervention, perhaps with reminders, but it was mentioned as an issue by only a small number of participants. There was also mention of laziness as a barrier to reduction, *"I want my applications to be there when I start my laptop up' which implies some laziness as well as efficiency"* (P58), *"I think it also comes down to laziness"* (P66). Researchers also identified a general lack of motivation on the part of students to reduce their consumption; *"at the moment there is no motivation to use less energy"* (P52), *"one of the main reason why they didn't monitor the current power usage was because they had not motivation or encouragement to do it"* (P95). Echoing a theme identified earlier, many researchers identified *Comfort and Convenience* as barriers to students behaviour change; *"participants are unwilling to change their hobbies just to save energy"* (P53), *"any energy intervention cannot interfere with their personal comfort"* (P98), *"convenience appears to also be a key factor in the student's choices"* (P83).

One striking theme that emerged consistently from the data was that of *Apathy*. Participants consistently reported a lack of interest in their use of electricity; *"Most of the interviewee's did not consider energy usage a concern of theirs"* (P23), *"they didn't think about they're electricity usage very much at all"* (P17),*"Even though I realise it is an issue, it has never been imposed on me culturally"* (P37), *"the topic of energy is quite boring*" (P89) *"it was a general lack of interest"* (P101), *"participants said that they did not actively think about their energy usage"* (P61),*"none of the people questioned, are actually interested in the information regarding their energy consumption"* (P47), *"participant one has no interest in the consequences of their actions"* (P52). Besides a general lack of interest in their use of electricity, there was also a specific lack of interest in reducing their consumption. *"I do not make a conscious effort to reduce my energy consumption"* (P81) *"He also admitted that this usage could be lowered but saw no benefit in lowering it"* (P30), *"YOLO"* (P8), *"not being entirely concerned about the effects of their usage resulted in them using more energy than necessary*" (P52), *"even though users may know they use too much energy they may not be willing to make changes"* (P45). These data outline a very difficult challenge for the designers of persuasive interventions. The prevailing opinion towards reducing energy consumption is a lack of interest and a lack of responsibility, reflecting the findings of [20]. The most striking aspect of these data is the honesty of responses, despite the fact that the attitudes expressed could be considered socially inappropriate.

Given all of the challenges identified to creating successful energy interventions for shared student accommodation, it is essential to now examine the suggestions provided by participants for potential solutions to this problem.

Category 4: Suggested Behavioural Solutions

These data (209 mentions) represent researchers' attempts to understand student's perceptions of interventions that may be effective in engaging students in reducing their energy consumption. Five distinct themes were identified in the data; *Lifestyle Changes, Rewards, Punishment, Competition, and Cooperation.* Some researchers identified *Lifestyle Changes* that participants expressed as acceptable; *"going out for football, gym, swimming or walking would cause in saving energy, because no one will use laptop, lights and so on"* (P2), *"I will turn off my light when I am sleeping and out of the room then I will not charge the* *laptop all day"* (P35). Suggestions such as these represent a tiny percentage of the total data, and given the evident lack of interest expressed by the majority of participants in saving electricity, they seem unlikely to be followed long term by a significant number of students.

Much more evident in the data was a wish from students for more obvious consequences of their energy consumption. Participants consistently identified the provision of *Rewards* as an attractive component of any intervention; *"a reward would make them motivated if involved with the saving of energy"* (P15) *"the need for students to feel that they are being rewarded for the time that they put into something"* (P84) *"being rewarded was a key part to engage the end-user"* (P85). Suggestions were also made of how a reward system could work in practice; *"A reward system whereby at the end of each month, the person who saved the most is rewarded"* (P102), *"rewards the user each time energy is saved when they manage to use it less"* (P6), *"A reward for the group or individual who has used least energy"* (P39). As a note of caution, however, one researcher identified, *"both of them were interested in the size of the reward in relation to the amount of effort they must carry out"* (P58). Some researchers identified the possibility of using virtual rewards, *"the app should include an achievement system similar to games consoles"* (P16), *"awarding points for a user reducing energy consumption which could then be used in-game"* (P33). However, the majority of participants who suggested the use of rewards as motivators preferred tangible incentives, *"Perhaps like, Amazon Vouchers for winners or something like that"* (P5), *"monetary rewards"* (P19), *"this reward was beer"* (P15), *"free gig tickets and downloads"* (P52), *"alcohol, vouchers, Facebook points"* (P59), *"£50 to everyone in the flat that saves the most energy…"* (P66), *"Err, free food, free drink, and money"* (P77). Interestingly, *Punishment* was identified by a small number of researchers as a useful means of motivating reductions in energy consumption, although the suggestions were quite vague; *"punishment would also encourage them to save energy"* (P87), *"I should create something social so that very wasteful people would be shamed in front of their friends"* (P101), *"making someone feel bad without restricting them physically or in reality is the way to go"* (P40).

The most commonly suggested means for motivating reductions in energy consumption was through *Competition*; *"the student demographic liked the idea of competing"* (P97) *"the best way of making people change their behaviour is to turn it into a competition"* (P16) *"The importance of competition is highlighted throughout all the sections of data collection"* (P42), *"giving the victory "energy bragging rights" was a phrased that was used in the focus group"* (P70). There were suggestions of how the competition could be facilitated in practice; *"maybe an inter-apartment competition"* (P10), *"groups competing against each other"* (P51), *"leader boards to publish your results to social networking sites"* (P79).

Researchers also found that participants suggested *Cooperation* as a technique to motivate students to reduce their energy consumption; "*It is clear that if some students started to save electricity it would encourage others as well… All participants agreed that such an environment where everyone is trying to consume less energy would push them to stop wasting it"* (P2), *"the behaviour of their flatmates influenced them, even against their own principles and highly held standards"* (P98), *"the importance of the application being inherently socially cohesive rather than competitive or divisive"* (P19), *"participants rejected individual competition in favour of a group-based activity"* (P98).

In summary, researchers have identified some very practical suggestions for motivating students to reduce their energy consumption. These all focus on providing consequences for behaviour, through either direct rewards or punishment, or through a larger social intervention involving competition and/or cooperation between students and reflect previous findings in the field (i.e. 17).

Category 5: Design Suggestions

These data (450 mentions) represent suggestions made by participants, and reflections by researchers, on the design of energy interventions. Seven distinct themes were identified in the data; *Platform, Simplicity, Visualisation, Functionality, Visual Appearance, Privacy Concerns*, and *Extravagant Requirements*. Most researchers gave some indication of the platform their participants indicated as appropriate for engaging with the intervention. Mobile devices were the most commonly mentioned platform, followed closely by laptops and personal computers; *"an app that can go on mobile phones"* (P14), "*they all wanted a mobile application"* (P38), *"it's got to be the laptop yep definitely"* (P3), *"phones, personal computers and laptops,"* (P29). Others emphasized the importance of an interface that is available across multiple platforms, thus allowing students to use their existing devices, *"it should be available multiplatform in order to allow access at any time"* (P83), "*The idea of using special device for observation of using energy was met negatively"* (P2), *"easily integrated with the gadgetry students use more regularly….. as opposed to brand new, stand-alone technology"* (P29). The last comment is interesting, since the established trend in household energy monitoring has largely focused on physical in-home devices. Researchers also identified that participants were attracted to *Simplicity* in design, and this was a consistently recurring theme in the data, *"simplicity appeared as a recurring theme in the design"* (P43), *"the users wanted a simplistic interface and didn't require a lot of time or effort"* (P85).

Researchers frequently reported on discussions of energy data *Visualisation.* They discussed the importance of using visualisations to help interpret the data for users*; "I don't think they will understand, if you show a lot of information"* (P28), *"would prefer it to be shown in a way they instantly understand"* (P31), *"the importance of clear, easy to read feedback"* (P73). Remarkably few suggestions were given for implementing those simple visualizations, but there was some discussion, *"simple bar graphs or pie charts"* (P88), *"The use of metaphors to represent the amount of energy saved"* (P42), *"having the object change colour"* (P31), *"No one remembers figures…..I'd remember the guy getting to the top of the steps"* (P77). Researchers identified that students often expressed an interest in seeing the consumption of individual appliances, *"ability to show the different electrical appliances so you can monitor them"* (P1), *"Comparing data on usage by device"* (P29), *"lots of detailed information about specific appliances"* (P53). Participants requested the ability to visualize expenditure, *"how much the energy is actually costing you",* (P5), *"giving a cash value to saving of energy"* (P7), to track progress over time, *"what would help them improve is seeing they're own 'self-improvement"* (P16), *"Recognition of progress"* (P20). Researchers identified that students valued the ability to see these data in real time, *"what people want to be able to do is walk around their flat and be able turn things off, and see the statistics change"* (P16), *"I would love a big board with lots of numbers moving around in real time. That'd be awesome"* (P77).

Researchers discussed the types of *Functionality* that participants reported as appealing. A number of researchers identified notifications and reminders as possible solutions, *"something which reminds you often"* (P18), *"A reminder to turn something off if you leave the room .... But without being annoying or nagging"* (P40), *"It could alert me half way through the day of how much energy I have used"* (P81). However, the dangers of such an approach were also discussed, *"the idea of a device telling someone to turn things off, was rather a troublesome subject.... If something explicitly tells me to do something I kind of resent it"* (P40). Researchers identified as important the integration of social features in any technological solution, *"should incorporate the major social networks into any design to encourage interactions between students"* (P83), "*a need for users to be able to communicate and share their results with others, for either competitive or cooperative/encouragement reasons"* (P45), *"those who are saving a lots friends will glorify them"* (P101). However, disadvantages of including social functions were also raised; *"Although this is a good idea one of the disadvantages could be that there may be rebellious students who want to boast how much energy they can use"* (P82). A number of researchers mentioned the wish for some form of automation or remote control as part of the intervention, *"one participant wants the app to be able to turn the heating on or off"* (P38), *"a push of a button can switch it all off"* (P14), *"a method of turning a light bulb off without them interacting with it"* (P87).

Interestingly, while participants seemed to want social features and interfaces that are accessible online across multiple platforms, they also expressed *Privacy Concerns*; *"the lack of privacy…was a great enough concern for the participants to not post sensitive data onto Facebook"* (P55), *"they wanted to be able to see friends' energy usage and vice-versa but not have anyone else be able to see their usage"* (P70), *"they would like to have a choice to share or not"* (P2). Interestingly, one researcher found that, *"it is important that students can activate and deactivate their energy tracking at any given time"* (P88). This concern, if implemented, would clearly undermine the accuracy and efficacy of any of the interventions based on feedback or providing consequences.

Researchers noted that participants often referred to the *Visual Appearance* of potential energy reduction technology. There was frequent that the final design, *"should look 'Cool'"* (P25), *"If the object is ugly it will not be used"* (P31),*"bright colours"* (P20). Indeed, many researchers expressed favour towards, *"a traffic light system, red meaning they are doing very bad, green meaning there are saving money and yellow being normal"* (P38). There were also quite a number of *Extravagant Requirements* expressed, ones which seem too vague too be realistically implemented given our current understanding of technology development; *"I think the summary of it all is it has to be something different each time to keep you interested….Monday it's this type of design. Every Tuesday it's something else"* (P55).

Design Implications

Given the breadth of challenges raised in these data, it seems unrealistic to think that a simple technology solution can have significant effects on the behaviour of students living in shared accommodation. The problem is clearly complex, which provides some explanation for the relative lack of success of technology-led energy interventions, which are typically designed based on simplified conceptions of behaviour. Any successful intervention must engage with and address precisely the types of issues raised in this paper.

Particularly challenging is the observed attitude that convenience, comfort and enjoyment seem to take precedence over any other concern, and that saving the environment is considered particularly lame (reflecting [1,17,20]). Moreover, there is a lack of willingness on the part of participants to accept responsibility for their own actions, as reported so succinctly by P52, *"participant one has no interest in the consequences of their actions."* The development of feedback technology is justified by the assertion that people are unaware of their own usage, and that feedback in itself should help people control and reduce their consumption [4,9]. This rather positive assumption was reflected by many participants and researchers in the data presented here. However, there was also a thread of critical reflection raised by a number of researchers, which cautioned against such optimism, *"one of the subjects said that even if they knew their usage it wouldn't stop them from being wasteful"* (P101).

Such complex human problems can perhaps be best addressed through human solutions. Participants have emphasized the highly social nature of the student lifestyle, and have expressed the importance to them of following the social norms and fitting in. Working with student groups to help establish a norm where this social pressure is exerted towards saving energy may be particularly effective, especially if this is in place from day 1 of students reaching their shared accommodation ([17] had similar suggestions). Interventions that integrate with real world campus activities to support and promote social interaction around themes of energy reduction may be effective (for example holding "no lights" parties), but must be careful to avoid being associated with the "lame" tag with which students typically refer to environmental concerns.

Crucially, our findings suggest that there is one relatively simple intervention that would have a significant effect on student's consumption of electricity in shared accommodation; the introduction of billed metering. Working with student accommodation providers to design technology to facilitate the introduction and administration of energy metering may be the area of most impact in this space. Importantly, students were generally positive about the introduction of billing. Inspiration could be drawn, for example, from China, where [13] reports that pre-paid energy has recently been introduced in some universities.

CONCLUSION

One hundred participants (students on an undergraduate HCI module) recruited approximately 300 further participants to carry out interviews and participatory design tasks, exploring the design of interventions to motivate reductions in energy consumption in shared student accommodation. Participants' and researchers experiences and reflections were analysed inductively to draw out clusters of themes in the data, and are presented here not to justify a theory, but to help designers maintain connection to participants (see [10]) in their design work.

The work goes beyond the qualitative environmental psychology research reported by [9] through its focus specifically on exploring the design of interventions that use modern technology such as smart meters, smart phones and social networking. It also builds upon the work of Toth et al., [20], which sought to understand the attitudes of teenagers towards energy consumption, but did not focus specifically on the context of student accommodation.

The research method employed principles of both participatory design and practitioner-based inquiry. Specifically, end users (i.e., students) were empowered to influence the design of technology that will directly affect them, through defining the very questions asked, carrying out the research and reporting the data themselves. Thus, the narrative generated from research is inherently and entirely composed of the practical lived experience and reflections of end users. Rather than having the assumptions of designers forced upon them, end users directed the course of the research. We feel this is a major advantage of the method employed.

Another advantage of the method lies in the honesty of the data gathered. It seems that the fact that peers rather than professional researchers conducted the research allowed participants to be uncommonly open and honest in their discussions. In addition, we feel that carrying out research in such close collaboration with students [15] is a uniquely productive approach to teaching and learning HCI at undergraduate level, and one capable of generating publications for tutors.

Ultimately, the contribution of the paper is in the very human reflections present in the data gathered by researchers, which will certainly inform our own project. However, we hope that the breadth and scope of experiences and reflections expressed in these data will encourage in other designers of energy interventions inspiration to engage with this challenge in a way that reflects the complex unpredictability and irrationality of the human experience, rather than appealing to theories that are not validated for this particular context. We hope this work will inspire novel and realistic types of energy intervention.

REFERENCES

1. Bates, O., Clear, A. K., Friday, A., Hazas, M., & Morley, J. Accounting for energy-reliant services within everyday life at home. In Proc Pervasive Computing (2012), 107-124.
2. Blythe, M., Wright, P., McCarthy, J., & Bertelsen, O. W. Theory and method for experience centered design. In CHI'06 extended abstracts (2006), 1691-1694.
3. Corbin, J., & Strauss, A. (Eds.). Basics of qualitative research: Techniques and procedures for developing grounded theory. Sage, 2008.
4. Darby, S. (2006). The effectiveness of feedback on energy consumption. A Review for DEFRA,486, 2006.
5. Erickson, T., Li, M., Kim, Y., Deshpande, A., Sahu, S., ... & Naphade, M. The Dubuque electricity portal: evaluation of a city-scale residential electricity consumption feedback system. In Proc ACM CHI (2013), 1203-1212.
6. Forlizzi, J., & Ford, S. The building blocks of experience: an early framework for interaction designers. In Proc ACM DIS (2000), 419-423.
7. Foster, D., Lawson, S., Blythe, M., & Cairns, P. Wattsup?: motivating reductions in domestic energy consumption using social networks. In Proc ACM NordiCHI (2010), 178-187.
8. Foster, D., Lawson, S., Wardman, J., Blythe, M., & Linehan, C. Watts in it for me?: design implications for implementing effective energy interventions in organisations. In Proc ACM CHI (2012), 2357-2366.
9. Froehlich, J., Findlater, L., & Landay, J. The design of eco-feedback technology. In Proc ACM CHI (2010), 1999-2008.
10. Gaver, W., Boucher, A., Pennington, S. & Walker, B. Cultural Probes and the value of uncertainty. Interactions, 6, 5 (2004), 53-56.
11. Ghassan, A., & Blythe, M. On legitimacy: designer as minor scientist. In Proc ACM CHI Extended Abstracts (2013), 2149-2158).
12. Gustafsson, A. and Gyllenswärd, M. The power-aware cord: energy awareness through ambient information display. Ext. Abstracts ACM CHI (2005), 1423-1426.
13. Liu, T., Ding, X., Lindtner, S., Lu, T., & Gu, N. The collective infrastructural work of electricity: exploring feedback in a prepaid university dorm in China. In Proc ACM Pervasive & Ubiquitous Comp. (2013), 295-304.
14. McWilliam, E. W(h)ither Practitioner Research?. The Australian Educational Researcher, 31 (2004), 113-126.
15. Neary, M., & Winn, J. (2009). The student as producer: reinventing the student experience in higher education. In: The future of higher education: policy, pedagogy and the student experience. Continuum, London, pp. 192-210.

16. Norman, D. A. Emotional design: Why we love (or hate) everyday things. Basic books, 2007.

17. Odom, W., Pierce, J., & Roedl, D. Social Incentive & Eco-Visualization Displays: Toward Persuading Greater Change in Dormitory Communities. In Workshop Proc. OZCHI (2008).

18. Olivier, P., and Wallace, J., Digital technologies and the emotional family. International Journal of Human-Computer Studies, 67 (2002), 204–214.

19. Rolfe, G. The theory-practice gap in nursing: from research-based practice to practitioner-based research. Journal of Advanced Nursing, 28 (1998), 672-679.

20. Toth, N., Little, L., Read, J. C., Fitton, D., & Horton, M. (2012). Understanding teen attitudes towards energy consumption. Journal of Environmental Psychology.

21. Vredenburg, K., Mao, J. Y., Smith, P. W., & Carey, T. (2002, April). A survey of user-centered design practice. In Proc ACM CHI (2002), 471-478.

22. Weiss, M., Mattern, F., Graml, T., Staake, T., & Fleisch, E. (2009, November). Handy feedback: Connecting smart meters with mobile phones. In Proc ACM MUM (2009), 15. ACM.

23. Wright, P., & McCarthy, J. Technology as experience. MIT Press, 2004.

24. Wright, P., & McCarthy, J. Experience-centered design: designers, users, and communities in dialogue. Synthesis Lectures on Human-Centered Informatics, 3(1), 1-123, 2010.

Measuring Social Media Activities of Dutch Museums: Developing a Social Media Monitor

Thijs Waardenburg, Rogier Brussee, Erik Hekman
HU University of Applied Sciences Utrecht
Postbus 8611 – 3503 RP Utrecht, the Netherlands
thijs.waardenburg@hu.nl, rogier.brussee@hu.nl, erik.hekman@hu.nl

ABSTRACT

In the research project "Museum Compass" we have developed a prototype of a social media monitor, which contains data of current and historic online activities on Facebook, Twitter, YouTube, Foursquare and Flickr of all registered Dutch museums. We discuss – mostly in a practical sense – our approach for developing the monitor and give a few examples result of its usage.

Author Keywords

Social media; analytics; metrics; cultural heritage; museums

ACM Classification Keywords

H.5.m. Information interfaces and presentation: Miscellaneous

INTRODUCTION

Social media have enabled easy and inexpensive interaction between millions of individuals and communities, and this has not gone unnoticed by the popular press, businesses and by scholars. Thus the online strategy of an organization has evolved from having a web site, to include a presence on social media. The question arises what all these social media activities actually bring. To answer that question, one has to start with measuring social media activities. However, current metrics solutions often consist of a confusing accumulation of statistics, across several systems, which reveal "little about online user behaviour, engagement and satisfaction" [3].

In this paper we discuss – mostly in a practical sense – the general approach and choices we made in developing a prototype of a social media monitor. The main goal of the museum monitor is to offer museum professionals and researchers better insight in the effects of their own social media usage and compare this with others in the cultural heritage sector. For researchers it gives the opportunity to consider communication within the sector as whole. The monitor, however, was developed in such a way that it could also be used for other sectors. In most of the paper, the word 'museums' and cultural heritage sector can therefore be substituted with companies in an industry, or a group of organisations that have some cooperation but also compete for a roughly similar audience.

The prototype was developed in the context of "Museum Compass"[1] a project that helps Dutch museums deploying 'crossmedial' strategies more effectively. In the project we developed several other services and tools to support Dutch museum professionals such as the "Museum Guide"[2], an online questionnaire that helps museums determining their current and desired cultural and societal identity (its 'archetype').

The organization of this paper is as follows. The next section is focused on related work, after that the approach and choices that we made are described. The conclusions and discussion are described in the remainder of this paper. A paper in preparation (part II) will focus on results and data-analysis.

RELATED WORK

Many articles, books, papers, etc. are written about social media, and various definitions are proposed. For instance, Kaplan and Haenlein [4] define social media as "a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of User Generated Content". Brussee and Hekman [2] propose a higher-level definition. According to them social media are "highly accessible media": social media characterized by the accessibility of the whole media supply chain to the general public. Where most definitions agree that current social media are digital by nature, they pose that the internet and web technologies are just very well suited for providing this accessibility.

The public and digital nature of social media activities allow for automatically monitoring and collecting data. Commonly used terms for such data collection are 'social monitoring', 'social media analytics', and 'social media metrics'. Many commercial off-the-shelf offerings can be found. They range from free 'one-size-fits-all' web 'dashboards', which collect data of one, or a limited number of platforms and present simple statistics, to extensive software that mines both structured and unstructured data and data from social and traditional media, and integrates quantitative data, qualitative data in the analysis often with

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

[1] http://www.museumkompas.nl (in Dutch, checked 20-01-2014)

[2] http://www.museumkompas.nl/museumwijzer/ (checked 20-01-2014).

accompanying consultancy writing reports. For an overview with over 230 solutions see http://wiki.kenburbary.com[3]. These solutions range from relatively simple and free of charge, to extensive and costly. Examples of social media reports are the Dutch social media monitor for healthcare ("Social Media Monitor Zorg")[4] and the British Culture24 report [3], titled "Let's Get Real: How to Evaluate Online Success?"[5].

According to Kaplan and Haenlein [4], integration is important because "what is true for different types of Social Media also holds for the relationship between Social Media and traditional media: Integration is key!" Murdough [5], on the other hand, stresses "the important rule is to focus on just a few metrics for each objective so that program evaluation remains simple and one does not end up in "analysis paralysis"". Bruns and Liang [1], state "for more sophisticated research programmes, and for the tracking and study of larger-scale datasets over longer time periods, more advanced and usually custom-made tools and methods are required." This is the approach we have taken.

APPROACH AND CHOICES

This research project is similar to, and partly inspired by the Culture24 project. An important difference is that the latter is based on social media data from *a selection* of museums, and *at a given moment*. The goal of the monitor that we are developing is to *continually track* social media activities of the '*whole*' museum sector, and mine its history. As mentioned in the previous section, there are many commercial off-the-shelf solutions available to monitor social media. However, we chose to develop our own because we believe that it offers better opportunities to experiment, customize, and learn. It allows us to directly access the detailed data required to get a better understanding of the subject by tracking individual posts, post-likes, retweets, etc.

Requirements

We identified three basic requirements for the monitor, which can also be considered as development phases:

Measuring social media activities of museums on a daily basis. To answer "How much effort do museums put in social media?"

Measuring the activities of the public related to a museum and its social media activity on a daily basis to answer "To what extend does 'the public' respond on the social media activities?"

Serving as a social media benchmark, so that museums can evaluate their activities and responses as well as compare themselves with their peers to answer "How do museums relate to each other, with respect to their social media activities?"

On the basis of popularity (or 'maturity') and type we chose the following five social media platforms: Facebook ('social networking'), Twitter ('micro blogging'), Flickr ('photo sharing'), YouTube ('video sharing'), and Foursquare ('location sharing').

Regarding the handling of data of social media platforms, we identified the following requirements for our monitor:

Data needs to be collected and stored in a database ('back-end').

Data needs to be interpreted and combined where possible.

Data needs to be presented ('front-end', in the form of a 'dashboard') for easy use by museum professionals and decision makers.

DATA COLLECTION

To collect data from any social media account, one needs to have their account identifiers (ID's) on the platforms they use. The official Netherlands Museum Register[6] does register the URL's of the websites maintained by museums, but does not register their account-ID's on social media platforms. We therefore needed to collect and verify this information ourselves. Collecting social media account-ID's is in essence a simple task, but it proved to be very time-consuming. Since it is difficult to automate the verification of authenticity of accounts, this was done manually. The protocol that we used to collect account-ID's was:

```
if account is linked on official museum
website with platform logo:

            accept account

else if search on social media platform
for at most three variants of name of
museum name returns account:

      if account has official museum
website as associated website:

             accept account

      else if  account comes up when
      googling museum name + platform
      name AND looks real by value
      judgement:

            accept account

      else: reject account
```

[3] http://wiki.kenburbary.com/social-meda-monitoring-wiki (checked 20-01- 2014).

[4] http://www.socialmediamonitorzorg.nl/ (in Dutch, checked 20-01-2014).

[5] http://weareculture24.org.uk/projects/action-research/how-to-evaluate-success-online/ (checked 20-01-2014).

[6] The Netherlands Museum Register, http://www.museumregisternederland.nl/ (in Dutch, checked 20-01-2014).

Rejected accounts were mostly pages automatically generated from Wikipedia and tourist information and easy to recognise. The rest almost all belonged to well-meaning individuals that write about large, well-known museums.

To collect the data we used the application programmers interfaces (API's) offered by the social media platforms. The earlier mentioned social media platforms all offer an API that enables basically anyone to develop an application that exchanges data with the platform. Using the scripting language PHP we extracted most data fields from the JavaScript Object Notation (JSON) returned by the platforms and stored the data fields in a MySQL database. We used PHP because it is widely used, supported by all platform API's, and because it made it easier to find freelance programmers. Wherever the platform allowed, we carefully kept track of individual actions of the organisations (posts, tweets, etc.) and the public (comments, likes, retweets) storing the ID's and creation timestamps provided by the platform and our own retrieval timestamps.

Each API (i.e., platform) has its specific set of data-elements. Which data-elements can be accessed depends on authentication levels. We distinguish three levels:

No authentication: 'Access token' and account-approval are not needed. Only basic account data can be received (e.g. account name, account-ID, profile image, etc.).

One-sided authentication. 'Access token' is needed from the platform provider, but account-approval is not needed. Publicly available / visible account-data can be received (e.g. messages, number of fans, etc.).

Two-sided authentication. Application needs to be registered at the platform; 'access token' and account-approval are needed. Extensive account-data can be received (e.g. friend list, private messages, etc.).

We generally focussed on the second level, as it allowed us to track the whole Dutch museum sector without having to make agreements with every individual museum. Fortunately, the data-elements that can be retrieved on the second level are fairly comprehensive. A Facebook 'post', for instance, is retrieved as a JSON document with 26 data-elements like 'id', 'from', 'to', 'message', 'likes', etc. next to the text of the post.

For practical reasons, we chose to make a selection of data-elements available from the platform, as it kept the size of the SQL schemas and queries in check. In addition some potentially available information required lots of additional queries to the platform. This is a problem as the number of requests per hour that can be made on a platform is fairly limited. For example, the Twitter (REST) API has a rate limit of 15 or 180 data requests per 15 minutes, depending on the type of data request. Flickr has a limit of 3,600 requests per hour. Facebook does not offer a clear insight in request limits at all but seems to be limited around 3,600 requests per hour per token, per IP-address. We found that data request limits are not always very 'strict', and can be apparently lowered for no obvious reasons. We dealt with this problem by daily scheduling the allowed number of data-requests by using 'cronjobs' until we had queried the activity of all museums. It would have been easier, but much costlier to simply buy the data[7]. This has the advantage that one actually receives *all* historical data: for instance, it turns out that the free API of Twitter 'only' provides 3,200 historical tweets per account. However only three museums were reported as having more than 3,200 tweets when we started and for financial reasons, buying data was not an option.

THE FRONT-END INTERFACE

After approximately one month, we collected enough data to make a number of basic data views for the front-end of the monitor, like tag-clouds, tables and graphs using Google Chart Tools[8] and other open-source tools (e.g. D3.js).

The front-end of the monitor is primarily intended for museum professionals. Data is presented from the point-of-view of a single museum, and museums can compare themselves with other Dutch Museums. For example, it allows them to view the change in the number of Facebook followers during a specific period and compare that with several other museums (e.g. of the same type or in the same city). This data is visualized in a graph, as in Figure 1.

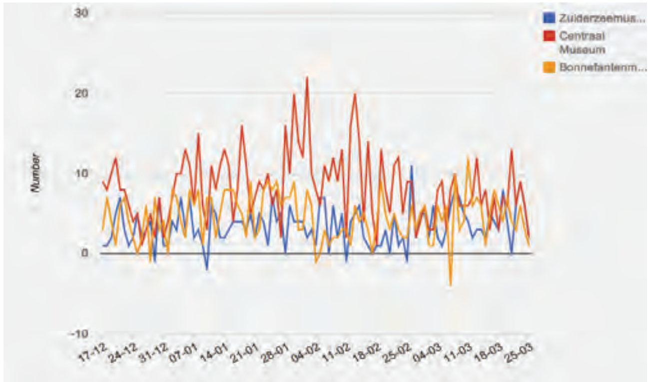

Figure 1: Screenshot of part of the interface for museum professionals

EXAMPLE RESULTS

A detailed data analysis will appear elsewhere ('Part II'). Here we give two examples of the kind of results that can be obtained by querying the database directly. This interface is only intended for researchers.

Sometimes museums react on other museums and thus show up as a 'public' reaction. Selecting these gives us a small social network in the sense of social network analysis (Figure 2), with avg. degree weighed by number of reactions = 4.85, and modularity = 0.605 representing the interactions of museums. It shows that museums are largely clustered by city (e.g. The Hague or Utrecht), and

[7] For instance: Twitter offers a service called 'Firehose' (accessible through 3rd parties like Gnip or DataSift).

[8] https://developers.google.com/chart/ (checked 20-01-2014).

secondarily by subsector (e.g. art museums or botanical gardens).

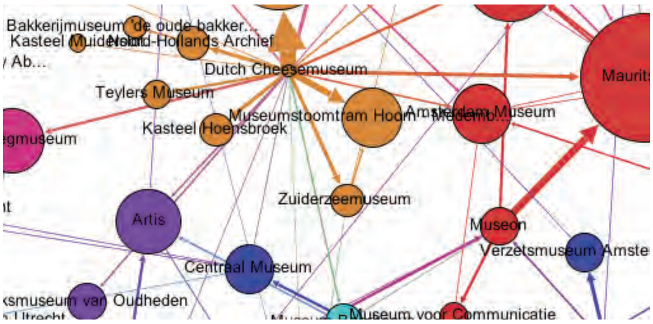

Figure 2: A fragment of the social network based on Facebook likes amongst Dutch museums.

Another example result is distribution over clock hours of Facebook posts of museums, and likewise, the distribution of the responses of the public. These are easily determined, since every Facebook post each post, comment and like etc. is tracked individually and is given a creation date stamp by the platform. This histogram (Figure 3) can help museums plan their Facebook posts for optimal impact.

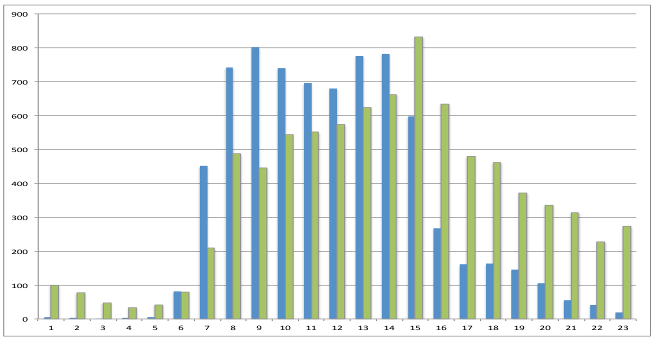

Figure 3: Histogram with the number of Facebook posts of Museums (blue) and 'the public' (green) over the 24 hours of the day. Museums post mostly during office hours peaking at 9h, the public peaks at 15h continuing during the evening and night.

DISCUSSION

We chose develop our own tool. However, we do not exclude that, in the future, other (commercial) solutions may also fulfil our requirements. The development of this tool has proved a useful way to get a better understanding of the subject.

The first version has a number of basic and separate data representations. We found that Facebook and Twitter provide by far the most interesting data and therefore intend to focus more on these platforms in the near future. This will also reduce the maintenance burden.

We chose to store only a *selection* of data, instead of collecting all available data. Long-term this could be a decision that we may come to regret as we may come to the conclusion that we need specific data-elements that we left out for analysis. Whereas rate limits remain a problem, a solution for the excessive schema problem is to collect and store all data in an alternative way, e.g., 'NoSQL' or 'NewSQL' database.

The protocol for verifying account-IDs should be refined however, as it is too subjective in its current form.

Ethical questions arise when we collecting social media data. We only collected public data that anyone can collect and the organisations that we tracked are public and visible organisations. However we will take care to anonymize data of individuals when publishing results.

CONCLUSION

We developed a prototype social media monitor to collect social media activities on Facebook Twitter, Flickr, YouTube and Foursquare of all Dutch registered museums. The monitor should be useful in other contexts as well. Practical considerations, in particular the use of platform API's and their limitations are discussed. We give an example of the results that can be obtained. Results of the data collection and analysis will be presented elsewhere.

ACKNOWLEDGMENTS

We thank the Foundation Innovation Alliance (SIA – Stichting Innovatie Alliantie) for funding RAAK-Museumkompas and all partnering museums.

REFERENCES

1. Bruns, A., Liang, Y.E.: Tools and Methods for Capturing Twitter Data during Natural Disasters. *First Monday 17*, 4 (2012).
2. Brussee, R., Hekman, E: Social Media are Highly Accessible Media. *Proc. of IADIS international conference on WWW/Internet 2009*.
3. Finnis, J., Chan, S., Clements, R.: Let's Get Real: How to Evaluate Online Success (2011). *Culture24*. http://weareculture24.org.uk/projects/action-research/how-to-evaluate-success-online/.
4. Kaplan, A.M., Haenlein, M.: Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons 53*, 1 (2009), 59-68.
5. Murdough, C.: Social Media Measurement: It's Not Impossible. *J. of Interactive Advertising 10*, 1 (2009), 94-99.

USER PERSPECTIVES

Dynamic Layering Graphical Elements For Graphical Password Schemes

Wouter van Eekelen
NHTV Breda University of Applied Sciences
Mgr. Hopmansstraat 1
4817JT Breda The Netherlands
wouter@picassopass.com

John van den Elst
NHTV Breda University of Applied Sciences
Mgr. Hopmansstraat 1
4817JT Breda The Netherlands
elst.j@nhtv.nl

Vassilis-Javed Khan
NHTV Breda University of Applied Sciences
Mgr. Hopmansstraat 1
4817JT Breda The Netherlands
khan.j@nhtv.nl

ABSTRACT

Based on a systematic review of 35 graphical password schemes, in this article a new classification and evaluation framework is proposed. When positioning existing schemes in this framework a novelty is discovered that wasn't previously described: a dynamic layered combination of graphical elements. Given this insight, a new graphical password scheme is created (*PicassoPass*). Positioned against other password systems, it has the potential to perform better on the combination of low memory burden and resistance to shoulder surfing attacks. A security analysis confirms its shoulder surfing resistance.

Author Keywords

Graphical password schemes; classification; evaluation; layering; *PicassoPass*; combination of graphical elements, shoulder surfing.

ACM Classification Keywords

H.5.2 [User Interfaces]: Evaluation/methodology; H.5.2 [User Interfaces]: Theory and methods; D.4.6 [Security and Protection]: Authentication; K.6.5 [Management of Computing and Information Systems]: Security and Protection.

INTRODUCTION

People are using passwords every day, multiple times; for online banking accounts, for social network profiles and to check their webmail from work. The great majority of all these digital systems have security measurements based on textual passwords. For over a decade the textual passwords' shortcomings have been documented [17]. One solution to these shortcomings is using graphical passwords [7, 13].

A definition of graphical passwords would be: `In a graphical password system, a user needs to choose memorable locations in an image. Choosing memorable locations depends on the nature of the image itself and the specific sequence of click locations` [20].

Scientists have looked into the possibility of graphical passwords, proposing numerous new ideas and systems [3]. Nevertheless, after all these years, despite the demonstrated benefits, graphical passwords have failed to replace textual passwords [6]. While textual passwords are mainly designed to serve technical goals first, graphical passwords are mainly designed to serve user goals first [10]. Such an approach has considerable advantages, but also raises challenges. Graphical passwords are more difficult to implement due to complex human factors that have to be considered [14].

Graphical password schemes can be grouped and differentiated within four different underlying ideas. As described by [14], [3] and [11], graphical password schemes are based on recall, recognition, cued recall or cued recognition. An alternative distinction between different types of graphical password schemes is `Cognometrics, Locimetrics and Drawmetrics' [1].

Although very useful, these classifications of graphical password schemes are limited when it comes to pertinent characteristics for design and development, such as security, technical aspects and graphical aspects. In this paper we expand aforementioned classifications of graphical passwords after having reviewed 35 password schemes. The contribution of this classification is an extensive evaluation matrix for designers and developers of graphical passwords. We strongly believe that such an evaluation matrix can help designers position their ideas and to serve as inspiration for novel solutions. To demonstrate this, we present an analysis resulting in a new graphical password solution which we call *PicassoPass*.

This paper is organized as follows. First we present the new classification and evaluation framework. Then we describe *PicassoPass*, a newly identified solution for graphical passwords. In order to contrast it against other systems, the advantages and shortcomings of *PicassoPass* are listed according to the new framework, and finally, the results of a study for shoulder surfing resistance are presented.

CLASSIFICATION AND EVALUATION OF GRAPHICAL PASSWORD SCHEMES

As described by [14], [3] and [11], graphical password schemes can be based on recall, recognition, cued recall or

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

cued recognition, or alternatively on `Cognometrics, Locimetrics and Drawmetrics' [1].

Recall is concerned with retrieving the correct answer from memory. Textual passwords work in the same way: the only thing users see is an empty input text field for which they need to recall the correct password. Most recall-based graphical password systems are the ones that require the user to draw something, usually on a grid [5]. Such a solution is also known as Drawmetrics [1].

In the case of cued recall, users have to find previously chosen spots from an image or picture [20]. An example of a cued recall based system is Pass-Go, which uses the game board structure of the game `Go'. By positioning the playing pieces, users can draw their password [15]. Cued recall is essentially a combination of recall and recognition.

Recognition based graphical systems work with a given image or collection of images (mostly displayed in a grid) at which users have to select (by recognizing) the correct spots or images, sometimes in a particular order [7, 9]. This is also called Cognometric [1]. Cued recall uses the given image but gives less cues as compared to recognition.

A final variant of graphical password systems found in literature, is Logimetric. As [1] describe it is `based on the method of loci, an old and well-known mnemonic`. The idea is that people have to retrieve objects from memory by mentally revisiting locations or stories. In a sense it is somewhat the same as cued recognition. An example of a Logimetric graphical password system is Story Scheme as mentioned by [11] `where the story or the semantic relationship between the images assists the user in the recognition of password images`.

We expanded above categories of graphical passwords into a new classification framework. Based on 40 different scientific publications (of which only a selection is referred to in this paper) covering 35 graphical password schemes, we identified a set of variables which can be used to classify and evaluate graphical password schemes. The used publications each describe one or more graphical password schemes. Some of them are the original, first papers on a specific scheme, while others reviewed similar schemes without mentioning a new scheme.

The information extracted from each publication was the name of the password scheme, a short description of the scheme and a list of its features. Given the identified variables and the 35 graphical password schemes we created a new classification framework. Regarding the password schemes and their particular characteristics, the complete set of variables identified in literature was aggregated into five main categories. Each category has one or more variables, on which a scheme was scored or for which was described how that scheme works. We do want to acknowledge that, although we did have a systematic process in categorizing the reviewed schemes, the process still remains to some extent subjective. Nevertheless, we strongly believe that such a classification can provide new perspectives to developers of new schemes and therefore want to contribute it to the community.

CATEGORIES AND VARIABLES

Our new classification and evaluation framework is presented in the form of a matrix comparing password schemes. The resulting matrix contains five `high level` categories: Memory, Complexity, Technical, Security and Graphical. Each category has multiple scaled variables, like password space for Complexity, combining graphics for Graphical, shoulder surfing resistant for Security, password storage for Technical and many more.

A simplified version of the evaluation matrix is presented in Table 1. This matrix contains all the identified categories and variables, but only shows the comparison between *PicassoPass* (which is introduced later in the paper) and a selection of other graphical password schemes. The main reason why those graphical passwords presented in this simplified matrix were selected, is that together they represent the full diversity of graphical passwords. The full list of all 35 graphical schemes that were evaluated is presented in Appendix A. Although all these schemes have been included in our investigation, due to space limitations only a selection of the matrix could be displayed. On request the authors can provide the complete matrix.

Memory

The first category is Memory, which is concerned with the underlying concept of a password scheme and how well it enables users to (easily) remember their passwords. For this purpose a password can be based on the variables mentioned earlier in this paper: recall, recognition, cued recall or cued recognition [3, 11, 14], as well as on Cognometrics, Locimetrics and Drawmetrics [1].

The second aspect of Memory is whether a user can create her/his own password or the password is generated by the scheme. User-created passwords are more likely to be remembered, while system generated passwords are often stronger and less likely to be guessed [3, 11]. A third possibility is that users, instead of creating their passwords, are selecting a password from a (proposed) collection provided by the scheme. This balances the forces between creating and generating passwords [2].

The impact on memory and the ability to remember a password is called memory burden: how much does a user have to remember so that she is able to input the password correctly in one attempt [3]. There are different techniques that can help to lower the memory burden [3, 7, 10], or limit the number of steps, of which the latter will also lower the strength of a password.

Using decoys that are (very) similar could have a negative effect on the memory burden, since users have to put more effort in remembering the correct image due to similarities and a higher chance that the wrong image is selected.

	Graphical Password System / Elements of (technical) design	Textual	PicassoPass	Pass Faces	Color Login	Déjà Vu	MARASIM	V-GO	GrIDsure	Pass Doodle	Patternlock	Inkblots	Story Scheme	Pass shapes
Memory	1: Recall 2: Recogn. 3:Cued Recall 4: Cued Recogn.	1	4	2	2	2	2	3	1	1	1	3	4	1
	1: Cognometrics 2: Locimetrics 3: Drawmetrics	-	2	1	1	1	2	2	1	3	3	1	1	3
	1: User chosen 2: User selected 3: System Gener	1 or 3	2	2	2	3	2 → 3	2	2	1	1	3 → 1	2	1
	Memory Burden	Length of min, 8	Mnemonic, for each step 1 elem.	Mult. sel. images	Mult. images of the same colors	Mult. gen. images	4 chosen images	Mult. elem. In a single image	4 connected loc. on a grid	Personal drawing	Drawing lines betw. 9 dots	Key-pairs of two for each inkblot	Mnemonic, for each step 4 pict.	Drawing limited to 8 stroke direc.
	Memory training and interference													
Complexity	Multiple steps with or without predefined order	Single step order	Multi step order	Multi step	Multi step	Multi step	Single step order	Multi step order	Single step order	Single step order	Single step order	Single step order	Multi step order	Single step order
	Password Space (complexity)	94^N	60^N	9^N			10000		10000		9!	52^N	3024	± 10000
Technical	Input of password (Touch/Click, keypad/keyboard)	Keys	Both	Click	Click	Click	Keys	Click + drag drop	Keys	Draw	Draw	Keys	Click	Draw
	Password Storage	Text	Digit			Text			Digit	Digit	Digit	Text		Text
	Multi platform	+/-	++	++	+/-	+/-	+/-	+/-	++		Mob.	++	+/-	
Security	Dictionary attack	- -		++	++	++	++	-	+	-	-	++	+	-
	Brute force attack	+/-					++				+			
	Shoulder surfing	- -	++	-	++	-	+	-	+	+	-	++	+	+
	Phishing attack	- -	+	+	+	+	-	- -	+	- -	- -	++	+	+
	Man in the middle attack	- -	++	+	-	-	++		++	-	-	- -	+	+
	Guessing attack	-	++	-	++	++	+	-	-	+	-	++	-	-
Graphical	Distinction colors, shapes /images	-	+	++	+	++	++	++	-	- -	+	+	+	-
	Combination of graphics	- -	++	- -	+	- -	-	+	- -	- -	- -	+	- -	- -

Table 1: Simplified Classification Matrix: - - is low score, ++ is high score

The last aspect of Memory is training. How many (training) trials does a user have to complete in order to successfully input the password within a reasonable time [11]. This variable is included in the classification matrix, but an actual comparison has been left out due to the lack of data about this variable for both *PicassoPass* and the other graphical password schemes.

Complexity

The second category is Complexity. Two variables of Complexity are identified: the actual, mathematical complexity [18] and the order of password input. The former is straightforward, although there are differences in the notation of the complexity.

Password input could be based on a predefined order or based on random input order [14]. When the scheme is order based, the user has to repeat the steps during input exactly as when the password was created. The more steps are required, the higher the chance that the user makes a mistake [3]. Other schemes are giving a user the possibility to input the password in any order, which can confuse the user due to forgotten steps or actions.

There is also a difference between inputting the password within a single environment, like a single window or input field, or within multiple environments like multiple windows or steps. Some of the schemes are combining single step input with multiple step input [18].

Technical

The Technical category is the third category of the classification matrix. The purpose of this category is to indicate how users have to enter their password, by means of keyboard, touch, point & click or by using gestures [3, 10]. Keyboard entry could be used in combination with displaying (alpha)numeric values on top of the graphical passwords that have to be inputted into a text field [4].

Another technical aspect of the schemes is how the password is stored. When it is stored as plain text it is less secure than a hashed or encrypted password [3]. Sometimes a \`portfolio\` of images is stored for each user [11], taking up a lot of disk space [10]. For only a few investigated schemes the (secure) storage of the password was described, so our comparison on this aspect is incomplete.

The last aspect of the Technical category is whether the scheme is suitable for various platforms, like mobile, ATM, computer, tablet and if this aspect was taken into account from the start [1, 3, 8].

Security

The next category is Security, which is reflected in how resistant a scheme is against different types of attacks. A dictionary attack is related to the uniqueness of a password and if applicable, how many hotspots are present within a graphical password scheme [18]. A hotspot is a popular spot within an image for many users, making it very likely that an average user has also chosen this spot. First trying to abuse these spots can limit the amount of work for attackers. A user-specific image collection (so no or as less as possible common data between users) narrows down the possibility of a dictionary attack [11].

When a scheme is resistant to brute force attacks [13] it is impossible to try all combinations, due to, for example, a time-out or variation within the scheme during login.

A very often discussed threat is shoulder surfing. Shoulder surfing is a capturing attack, in which someone tries to look over the shoulder to capture the password [3]. This can also be achieved with recording devices like camera's [3], but also by using keyloggers, screenscrapers (to see what is happening on screen) and mouseloggers [18, 12].

A technique to counter shoulder surfing is the use of decoys [8] so malicious users are confused or cannot detect the correct answer unless they capture multiple trials of the login sequence.

A fishing (also known as phishing) attack stands or falls on how well users can describe their passwords [3, 14]. Some schemes are using randomized graphics which are highly recognizable for users while they are very difficult to describe due to not having any resemblance with everyday objects.

Another attack is the man-in-the-middle attack which is based on capturing the data transfer between user and the validating system [3]. Sometimes the actual password can be captured because it is sent as plain text. Sending it in an encrypted format is not always more secure, an attacker could hijack an encrypted password and use this to login when there is not a validation of when the password is inputted or from which device or location it was submitted. The encrypted password is likely to be the same every time, unless a time-based or one time valid token or randomized data is being used [11].

The last attack being defined within the Security category is guessing attack, where personal information, like gender or individual preferences, is used to guess the password. An example of information that can be used for a guessing attack is that people prefer faces of the opposite gender [3, 14].

Graphical

The final category is Graphical, with two variables: distinction and combination. The better the distinction is between colors, shapes and images, the fewer mistakes users make during input [1, 8], however it also increases the risk for successful dictionary and shoulder surfing attacks. Note that a good distinction between shapes also makes it possible for colorblind people to use the password scheme [3].

Combining graphical elements increases the strength of passwords and the theoretical password space, but is being used in only a very few graphical password schemes. Typically the used images or graphical elements that are being displayed are positioned within a grid [3, 14] instead of being combined. This aspect is discussed in detail in the next chapter.

DYNAMIC COMBINATION OF GRAPHICAL ELEMENTS

Only a few graphical password schemes are combining graphical elements. The main reason is likely that combining them increases complexity for the user, although most of the time it also increases the strength of the password and the theoretical password space, thus enhancing security.

A well known example of a graphical password scheme is Passfaces [3, 10, 15, 19, 21]. Passfaces only displays nine different images, limiting the password space to 9^N (N is the number of password images). If graphical elements are combined, especially in a dynamic manner, the password space could be extended.

An example of a graphical password scheme that uses this technique is Picture Password [10, 15]. Users can select two images from a grid, which acts similar to the shift key on a keyboard and forms a unique combination. However, Picture Password increases it's password space mostly by displaying 30 different graphical elements. Fitts' Law from 1954 then comes in: the time to point to a target depends on the distance (or total size) and size of the target. So, when the distance (or total size) becomes larger and the targets smaller, the performance becomes slower [20].

The best approach would be giving users a limited amount of clickable choices, while at the same time they have more possibilities. Layering could provide such approach: an image is constructed out of different layers. One layer could be for example a shape and another layer the color, as illustrated in figure 1. If a generated image has for example 12 different clickable choices and for each choice a shape and a color are combined, then it would result in a password space of 24 for a single image.

A color and a shape are two different things that humans can distinguish. So it doesn't matter that they are combined, they could also be presented uncombined so the image would have 24 different clickable choices with only a shape or color. When users know they need to select the correct color, they can mentally filter the other information and ignore what they don't need, like shapes.

If someone would look over the shoulder, she/he only sees that the user selects for example a red star. But was it selected because of the color red or because it was a star shape? Adding more layers will complicated things more for shoulder surfers, especially if the combination of layers is different every time (dynamic).

PICASSOPASS

The aforementioned classification matrix and the resulting observations regarding dynamically combining graphical elements, served as inspiration for a novel scheme. This scheme, called *PicassoPass*, scores very high on a particular variable which was underrepresented in existing systems: combining graphics. *PicassoPass* is a challenge-response based graphical password system. It dynamically combines graphical elements in different layers, which hasn't been described previously.

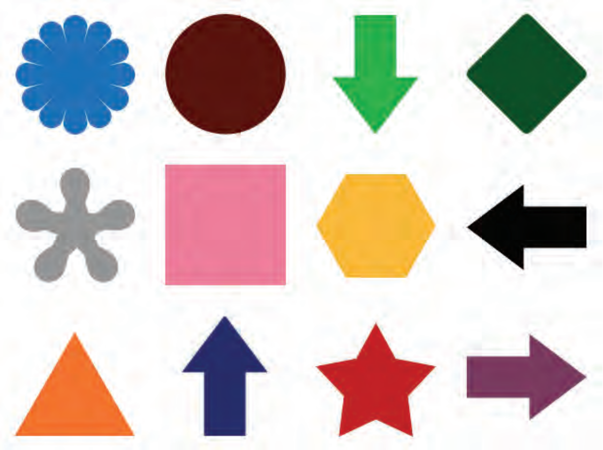

Figure 1: Two layers (shape and color) in one image

During login, *PicassoPass* presents a sequence of grid-based images. This is called a `challenge`. The task for the user is to select the correct cell from the grid at each step. What the correct cell is, depends on what the user has chosen as correct when creating the password.

Graphical

In *PicassoPass* each cell is a (random) layered combination of four different things: a basic shape (for example square or triangle), a color, a character from the alphabet and a shape based on a theme. This is presented in figure 2 and 3.

Instead of presenting a grid of 60 elements, the layering makes it possible to display a grid with only 12 elements, which needs less space on screen and at the same time inhibits shoulder-surfing. With every login the elements are randomly combined. When a user logs in, an attacker would not know why the user has selected a cell, since there are five different possible reasons (the four mentioned earlier, together with the position of the correct cell in the grid, see figure 2). It would require multiple captures of the login process to rule out all potential reasons.

For every grid / step, the user has chosen what selector is used, like the shape, character, color or the position of the cell. The user is going through each grid one by one until she/he has finished the challenge manually. An example could be that with the first grid, red is correct, the second top left position and at the third grid the circle is correct and the user finishes the challenge.

Complexity

The theoretical password space of *PicassoPass* is higher than (four digit) PIN-based password systems, yet lower than textual passwords with a length of five alphanumeric characters. For each grid, there are 12 distinct locations with each cell having four different elements combined: color, shape, theme and an alphabetic character. So, the possible combinations of each individual grid is 12x5 = 60. If the graphical password has four grids, it would be 60^4, or 12,960,000 possible combinations. A PIN of four digits has (10^4) 10,000 possibilities while a textual password with five alphanumeric characters including upper- and lowercase and symbols has (94^5) 7,339,040,224 possibilities, which is an enormous difference with the four digit PIN code.

Figure 2: Different Layers of PicassoPass

Technical

PicassoPass can be used on multiple platforms, since the 60 different elements are positioned within a grid of 12 elements.

A prototype was made for mobile devices with a small screen resolution of 320 pixels width and 450 pixels height and the elements (including the theme shapes) are still distinctive enough.

ATMs with touchscreens could also profit from *PicassoPass*, since due to the layering and random combination of elements, capture attacks will not work unless the target has been recording multiple times with drawing cash during the period the ATM was altered. When ATMs do not have a touchscreen, the amount of cells could be limited to nine so the position of cells corresponds with the actual keyboard layout of the keypad of the ATM. Web based solutions could also use *PicassoPass*, since point & click will also work with the grids and can be scaled up to be used on screens with larger resolutions.

Memory

PicassoPass uses the combination of graphical elements for a mnemonic approach [11]: a story assists the user in the recognition of graphical elements (cued recognition).

A positive effect of a story approach is that it contributes to a better recalling of a password: when the items or objects that need to be remembered can be associated with something concrete [22], they will be easier to remember [20]. This especially applies to semantically meaningful content like concrete images or real-world scenes (as described by Norman in 1988) [20], which are easier to remember than abstract images.

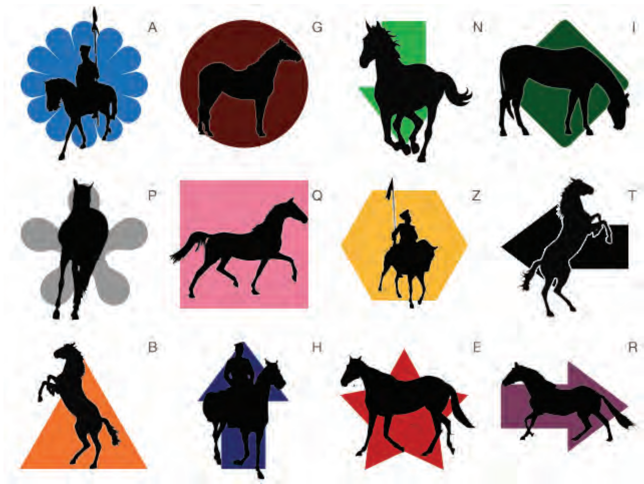

Figure 3: The final result of dynamically combined layers

The storage of the image in the long-term memory is not based on storing the actual image itself, but instead a `meaningful interpretation' as described in 1977 by Mandler & Ritchey [19, 20]. At the same time, there is a preference for images that are symmetric so memory load can be reduced [13, 16].

To aid users of *PicassoPass* with remembering their password, the theme shapes can be used to create a mnemonic story. An example could be `the blue horse jumped over the green car`. Every underlined word could potentially be a grid / step. To make the above example even stronger, it could be appended with `that has a yellow star on top`.

This story-based approach of *PicassoPass* with clear and distinct colors, shapes and images, gives it a significant advantage for memory burden and reducing mistakes.

Due to the mnemonic and graphical approach, users that are illiterate can still login. The layering is also beneficial for color-blind users that have problems distinguishing colors: they can use the other elements like position, shapes, themes and characters.

When the user has forgotten the password, a new one can be requested by entering their username and email address. An email with an unique URL with limited lifespan (1 hour maximum) will be sent to the user, after the username and email address are validated against the list of registered users. When the user clicked the one-time valid URL, the site requests to enter the same username and email address to verify the user. If the verification is successful, the user can create a new password.

Security study: comparison of shoulder surfing attack of PicassoPass with current password schemes on a tablet device

To test resistance for shoulder surfing an online survey was conducted. 57 participants responded out of 120 sent invitations. The only requirement for participation was perfect (or corrected) vision. No additional demographic information was recorded.

Each participant was shown one video of someone entering a password on a tablet device, filmed as the viewer was watching over the shoulder. Participants were divided into three groups. For each group, the used password technique was different. One group of participants saw a numeric password, another group saw a gesture and a third group saw *PicassoPass* (see figure 4). Participants were then asked: "What was the password the user inserted?". After viewing the video, the participant had to select the correct answer from a set of six possibilities.

For example, in the case of the numeric password, the video depicted a user tapping the "2998" numeric code to unlock the tablet. After viewing this short video, participants were asked the question: "What was the password the user inserted?". Participants got the following six options to choose from: "0987", "1234", "8463", "2998", "2292", "3015". These options were randomly ordered for each participant.

Similarly, in the case of the gesture password, participants, after viewing the video with the user unlocking the tablet with a gesture, were asked the question: "What was the password the user inserted?". Participants then saw six images each depicting a possible gesture with the help of an arrow-line.

Finally, in the case of *PicassoPass*, the video depicted a user going through three screens of *PicassoPass* to unlock the tablet. Then, participants got six options of possible element combinations to choose from.

Our null hypothesis of the aforementioned setup is: H0: “There is no difference between the three password techniques when it comes to shoulder surfing attacks.”

In total there were 57 participants (numeric: 18, gesture: 17, *PicassoPass*: 22). Table 1 shows the survey results for the number of successful and unsuccessful participants in guessing the passwords for the different password methods.

The two variables were: v1: password technique, v2: shoulder surfing attack. Both of them are nominal with possible values: v1=[Numeric, Gesture, PicassoPass] and v2=[successful, unsuccessful]. Since both of the variables are nominal, the statistical test needed to test the hypothesis is chi-square [7]. Thus, the value of chi-square was 40,94 which was significant at the .1% level with 2 degrees of freedom. That means the H0 can be rejected. By having a look at the contingency table it is clear that *PicassoPass* is significantly superior to the two existing password insertion methods, when it comes to resistance to shoulder surfing attacks.

The results of this between-subject study design show that none of the 22 participants who were assigned to *PicassoPass* correctly guessed the password, while almost everybody correctly guessed the numeric password (see table 2). This confirms the potential of *PicassoPass* to protect against shoulder surfing attacks.

Figure 4: Stills of the three different videos

	Shoulder attack	
Interface	Successful	Unsuccessful
Numeric	17	1
Gesture	13	4
PicassoPass	0	22

Table 2: Number of successful and unsuccessful participants in guessing the passwords for the three password methods

FUTURE RESEARCH

Although these promising first results could be an indication that *PicassoPass* has potential to be an adequate graphical password system, a more complete investigation is needed.

Future studies on the other main categories in the proposed classification framework should confirm this, in particular regarding the expected performance on memory burden and recall. In this section we elaborate three directions we are interested in exploring in the future.

Memory burden

PicassoPass is expected to perform especially well on memory burden (due to its story based approach with clear and distinct colors, shapes and images) and on shoulder surfing resistance. The latter has been tested and confirmed. Future studies should investigate to what extent it enables users to remember their password and clarify issues related to training and interference (for example if using multiple password could lead to password interference).

Usability testing

In this paper we presented a study based on finger-based interaction on a tablet device. Yet we envision PicassoPass to be generically used in all sorts of devices. ATM machines also imply touch-based interaction but what about desktop computers or smart TVs? On such devices mice, keyboards and remote controls are the primary means of interaction. We are interested in investigating how well does PicassoPass performs in terms of efficiency but also satisfaction from the user's point of view.

User Customization

The version we present in this paper has a set of icons showing sports (figure 2) and horses (figure 3). Nevertheless, this set of icons could potentially be of any given theme. One could imagine being able to customize the icon set based on a favorite movie or TV show or video-game. That could of course work for a specific set of devices but at the same time raises new, interesting questions of technical nature and its generic applicability. The similarities of the icon set could also lead to interference of remembering the password due to too many similarities and thus making it harder for the user to correctly input their password.

Security threats

Although shoulder surfing resistance is confirmed in tests, repeatedly shoulder surfing on the same user could possibly result in a successful guess of the password. Other security threats are also noteworthy to investigate, especially how well *PicassoPass* stands against dictionary and brute force attacks.

CONCLUSION

A systematic review of 35 graphical password schemes yielded a new classification which we propose as a novel way of looking at password schemes that would help designers of such schemes position their work. This classification allowed us to identify that combining graphical elements for a graphical password scheme is not often utilized, let alone a dynamic combination of layers which increases password space.

Based on this discovery, we developed a new graphical password scheme called *PicassoPass*. *PicassoPass* is a challenge-response based graphical password system that uses cued recognition. Its novelty is that it dynamically combines graphical elements in different layers. Its resistance to shoulder surfing attacks has been tested and confirmed. These results proved that it has potential to be investigated further.

REFERENCES

1. A. D. Angeli, L. Coventry, G. Johnson, and K. Renaud. Is a picture really worth a thousand words? exploring the feasibility of graphical authentication systems. International Journal of Human-Computer Studies, 2005.
2. K. Bicakci, M. Yuceel, B. Erdeniz, H. Gurbaslar, and N. Atalay. Graphical passwords as browser extension: implementation and usability study. Technical report, 2009.
3. R. Biddle, S. Chiasson, and P. van Oorschot. Graphical passwords: learning from the first twelve years. Technical report, 2011.
4. S. Brostoff, P. Inglesant, and M. Sasse. Evaluating the usability and security of a graphical one- time pin system. In Proceedings of the 24th BCS Conference on Human Computer Interaction - HCI2010, 2010.
5. K. Chalkias, A. Alexiadis, and G. Stephanides. A multi-grid graphical password scheme, 2008.
6. S. Chiasson and et al. A second look at the usability of click-based graphical passwords. In ACM SOUPS, pages 1-12. Press, 2007.
7. S. Chiasson, P. C. V. Oorschot, and R. Biddle. Graphical password authentication using cued click-points. In 12 th European Symposium On Research In Computer Security (ESORICS), 2007. Springer-Verlag, 2007.
8. A. De Angeli, L. Coventry, G. Johnson, and K. Renaud. Evaluating the usability and security of a graphical one-

time pin system. In International Journal of Human-Computer Studies, volume 63, 2005.

9. A. E. Dirik, N. Memon, and J. camille Birget. Modeling user choice in the passpoints graphical password scheme. In In SOUPS 07: Proceedings of the 3rd symposium on Usable privacy and security, pages 20-28. ACM, 2007.
10. S. Gkarai and A. Economides. Comparing the proof by knowledge authentication techniques. In International Journal of Computer Science and Security, volume 4, pages 237-255, 2010.
11. R. A. Khot, K. Srinathan, and P. Kumaraguru. Marasim: a novel jigsaw based authentication scheme using tagging. In Proceedings of the 2011 annual conference on Human factors in computing systems, CHI '11, pages 2605-2614, New York, NY, USA, 2011. ACM.
12. M. Kumar, T. Garfinkel, D. Boneh, and T. Winograd. Reducing shoulder-surfing by using gaze-based password entry, 2007.
13. D. Nali and J. Thorpe. Analyzing user choice in graphical passwords. Technical report, 2004.
14. X. Suo, U. Direction, Y. Zhu, X. Suo, and X. Suo. A design and analysis of graphical password, 2006.
15. H. Tao. Pass-go, a new graphical password scheme, 2006.
16. J. Thorpe and P. van Oorschot. Graphical dictionaries and the memorable space of graphical passwords. In Proceedings of the 13th USENIX Security Symposium, 2004.
17. J. Thorpe and P. van Oorschot. Towards secure design choices for implementing graphical passwords, 2004.
18. J. Thorpe and P. van Oorschot. Human-seeded attacks and exploiting hot-spots in graphical passwords. In Proceedings of the 16th USENIX Security Symposium, 2007.
19. S. Wiedenbeck, J. Waters, A. Brodskiy, and N. Memon. Authentication using graphical passwords: Effects of tolerance and image choice. In In First Symposium on Usable Privacy and Security (SOUPS 2005, pages 1-12. ACM Press, 2005.
20. S. Wiedenbeck, J. Waters, J. camille Birget, A. Brodskiy, and N. Memon. Passpoints: Design and longitudinal evaluation of a graphical password system, 2005.
21. S. Wiedenbeck, J. Waters, L. Sobrado, and J. Birgit. Design and evaluation of a shoulder-surfing resistant graphical password scheme. In Proceedings of the advanced visual interfaces symposium, 2006.
22. L. Zin, Q. Sun, and D. Giusto. An association-based graphical password design resistant to shoulder-surfing attack. In Proceedings of the IEEE international conference on multimedia and expo, pages 245-248, 2005.

APPENDIX

A. List of Evaluated Graphical Password Schemes

The following password schemes were evaluated using the comparison matrix, in random order. Since *PicassoPass* is based on the results of the comparison matrix, and textual passwords not being a graphical password scheme, the total number of schemes is 35.

- Awase-E
- Passfaces
- ColorLogin|
- Déjà Vu
- GPEX
- GPI and GPIS
- MARASIM
- Picture Password
- Use Your Illusion
- V-GO
- VisKey
- VIP
- DAS
- GrIDsure
- PassDoodle
- Pass-Go
- PassShapes
- PatternLock
- RAF
- CCP and PCCP
- ImageShield
- Inkblots
- Jiminy
- Loci-based
- PassPoints and PassPoints blur
- Story Scheme
- ColorPIN
- Color-rings
- Gaze based
- Movable frame
- S3PAS
- ShieldPIN
- SlotPIN
- CuePIN
- Convex Hul

Chi sparks 2014

Persuasive by Design: a model and toolkit for designing evidence-based interventions

Sander Hermsen
Utrecht University of Applied Sciences
Padualaan 99, 3584 CH Utrecht, Netherlands
sander.hermsen@hu.nl

Reint Jan Renes
Utrecht University of Applied Sciences
Padualaan 99, 3584 CH Utrecht, Netherlands
reintjan.renes@hu.nl

Jeana Frost
Free University Amsterdam
De Boelelaan 1081, 1081 HV Amsterdam, Netherlands
j.h.frost@vu.nl

ABSTRACT

Despite increased interest in applying psychological theory to the practice of designing behavioral change interventions, design professionals often lack adequate knowledge and resources to do so. In this paper, we present a tool to help professionals in the creative industries design evidence-based health interventions, the Persuasive by Design model. This paper describes the contents and application of the model as well as plans for further development and testing.

Author Keywords

Model; design tool; persuasion; design; evidence-based; interventions; persuasive technology.

ACM Classification Keywords

H.1.2 User/Machine Systems: Human factors

INTRODUCTION

Recent design research literature indicates an increased interest in applying insights from psychology and related sciences to design behavioral interventions. This interest spans the fields of sustainability (e.g. [1]), health (e.g. [2]) and mobility (e.g. [3]). So called "evidence-based" interventions have been shown to be both more effective at changing behavior [4] and to result in great increases in the decisional accountability of the designer [5].

Despite interest in applying current psychological theory to design practice, a disconnect remains between the fields of design research and service design on the one hand, and (cognitive) psychology on the other. Designers often view cognitive psychology research as "impenetrable" [6]. The psychological theories and models in current use within design suffer from limitations. Existing theories, such as Theory of Planned Behavior [TPA, 7], the Health Belief Model [HBM, 8] and the Fogg Model [9] do not address all aspects of behavior, and offer a limited view of persuasive interventions. These shortcomings may severely reduce the potential efficacy of any designed intervention based on these models. In this project, we propose a behavior change model to inform the design of evidence-based interventions. In this paper, we discuss the model, its practical application and plans for further development and testing.

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

PERSUASIVE BY DESIGN MODEL

The model includes both the contexts and intervention strategies for the intervention, displaying them in a set of color-coded layers and suggests a series of questions to help the designer address each aspect of the target behavior. The three contextual layers include: a blue layer with two different modes of behavior (reflective and automatic), a red layer that displays biases and other internal threats for behavior change, and a green layer displaying the social influences on our behavior. The two intervention layers include: a black layer that displays methods to change reflective behavior and a purple layer with strategies to target automatic, reflexive behaviors.

The proposed model improves upon existing tools to inform the design process in a few ways. Firstly, the model offers a broader range of possibilities for persuasive intervention design than TPA, HBM or the Fogg Model. Secondly, the model reflects current thinking on behavior. It takes into account the observed dichotomy between reflective and reflexive behaviors, and explicitly incorporates social influences and cognitive biases. Lastly, the layout and presentation of the model is designed for intuitive use.

DEVELOPING PERSUASIVE INTERVENTIONS

To introduce the model to creative industry professionals, we developed a workshop in which the model is presented. In this workshop, we introduce the elements of the model and the sets of questions associated with each one. Participants use the model and the questions to analyze target behaviors and select an appropriate strategy for the design of persuasive interventions.

Reflexive and reflective modes of behavior (blue layer)

The foundation of the model (fig. 1, blue layer) reflects the notion that most of our behaviors are executed in one of two modes: either automatically or with reflection.

The reflexive, automatic system acts upon the (often unconscious) perception of a cue, which turns on a habitual routine that leads to behavior. This system is fast and efficient, but because its execution is based upon prior experience without adaptation to the current situation, it is not always effective.

Reflective behavior, on the other hand, is best viewed as a self-regulation cycle reminiscent of a thermostat [e.g. 5, 12]. We compare our goals to current behavior. Upon noting a discrepancy, given enough motivation, ability and opportunity, we change our behavior, monitor our changed behavior, compare once again our current behavior to our goals and so on, until our goal is reached. This reflective mode is cognitively costly in that it requires conscious effort.

Viewing this foundational layer, the blue layer, triggers designers to think about the qualities of the target behavior or – in most cases – the chain of target behaviors. For instance, to develop an intervention to reduce shower length and save water, the question set derived from this part of the model suggests how developers should differentiate between an initial phase in which adaptation of the intervention is the goal, followed by an implementation phase where the actual behavior change takes place. For each separate link in the target behavior chain, the question sets and the model[1] inform the design of the intervention.

Of course, many of our behaviors are complex and engage both systems [10]. A set of questions in the model focuses designers on what aspects of the target behaviors are automatic in nature, e.g. habits and impulses, and which aspects are reflective.

The model also presents users with opportunities for increasing reflection about otherwise automatic processes. Behavioral research suggests that automatic behavior such as taking a shower every morning can be influenced by offering new cues or hiding undesired ones, for instance by setting favorable defaults [11]. When this is impossible or not ethically viable [12], automatic behaviors can be interrupted for reflective change, e.g. by giving feedback on habitual behavior. The purple layer of the model illustrates the possibilities for intervening in automatic behavior. In the case of the shower intervention, a visible cue can make the time spent in the shower tangible, thereby disrupting habitual behavior.

To sustainably change reflective behavior, the model suggests keeping goals and norms salient, offering feedback on current behavior and providing action plans [13], as displayed in the model's black layer. Take the case of how to design an intervention to reduce home heating energy use by promoting roof insulation. The black layer in the model indicates the utility of social norms to change behavior. Designers, viewing the model, might then create a searchable color-coded heat map visible at the level of individual homes.

Threats to reflective behavior change (red layer)

Everyday experience teaches us behavior change is not as easy as simply following the proposed self-regulation cycle until we hit upon success. The red layer in the model displays the many threats to behavior change that may occur. For each target behavior, possible deficiencies, biases and other threats to behavior change can be considered by answering corresponding questions. The red layer and the accompanying questions enable designers to reflect on these threats. Is the way the intervention is set up prone to induce resistance in the target group? Is the target group capable of judging their own behavior? In the example of an intervention to reduce energy consumption, the barrier to change behavior may be awareness. The heat map would make the consequences of poor roof insulation visible to the consumer.

A further group of questions included in the model addresses motivation, ability and opportunity to implement new behavior. These may provide barriers that prevent behavior change and are more often than not the weakest link in a designed intervention. In the heat map example, intervention designers may through answering these questions realize that even though the heat map makes it possible to experience roof heat loss, which can provide clear motivation to save both money and energy, perceived barriers such as cost and effort may seem too big to make behavior change possible. The black layer of the model suggests the intervention should offer an action plan fitting to the feedback. To remove felt barriers, designers may decide to include attic clearing services for a minimal extra fee in the intervention or display direct connections to available funding.

Social influences (green layer)

An effective model of behavior change should also take into account the fact that humans are social beings, not autonomous entities oblivious to social influence [14]. The green layer in the model reflects social influences on the reflexive cycle.

To attend to social influences, the model includes a set of questions about the influence of social processes in all their complexity. While social comparison can be a powerful motivator to change, finding out you are doing better than your peers can impede performance. Similarly, social commitment and peer pressure may enhance motivation to take part in energy saving measures, but social validation – everybody has this problem, so why should I bother to change this – may decrease said motivation. When designing an intervention for example, to save energy, these social processes need to be taken into account. The green layer in the model and the answers to the accompanying questions enable designers to build interventions that benefit from social influences and avoid undesirable effects.

DEVELOPMENT AND TESTING

A first version of the model was based on an extensive review of recent behavior change literature, followed by two co-design sessions in which we introduced drafts of the model to professionals from the creative industry. These sessions enabled us to do some fine-tuning and led to the conclusion that because of the inherent complexity and nuances of this model, we also need to craft a proper introduction and background explanation.

[1] A full set of questions accompanying the model is available from http://www.touchpoints-hu.nl/

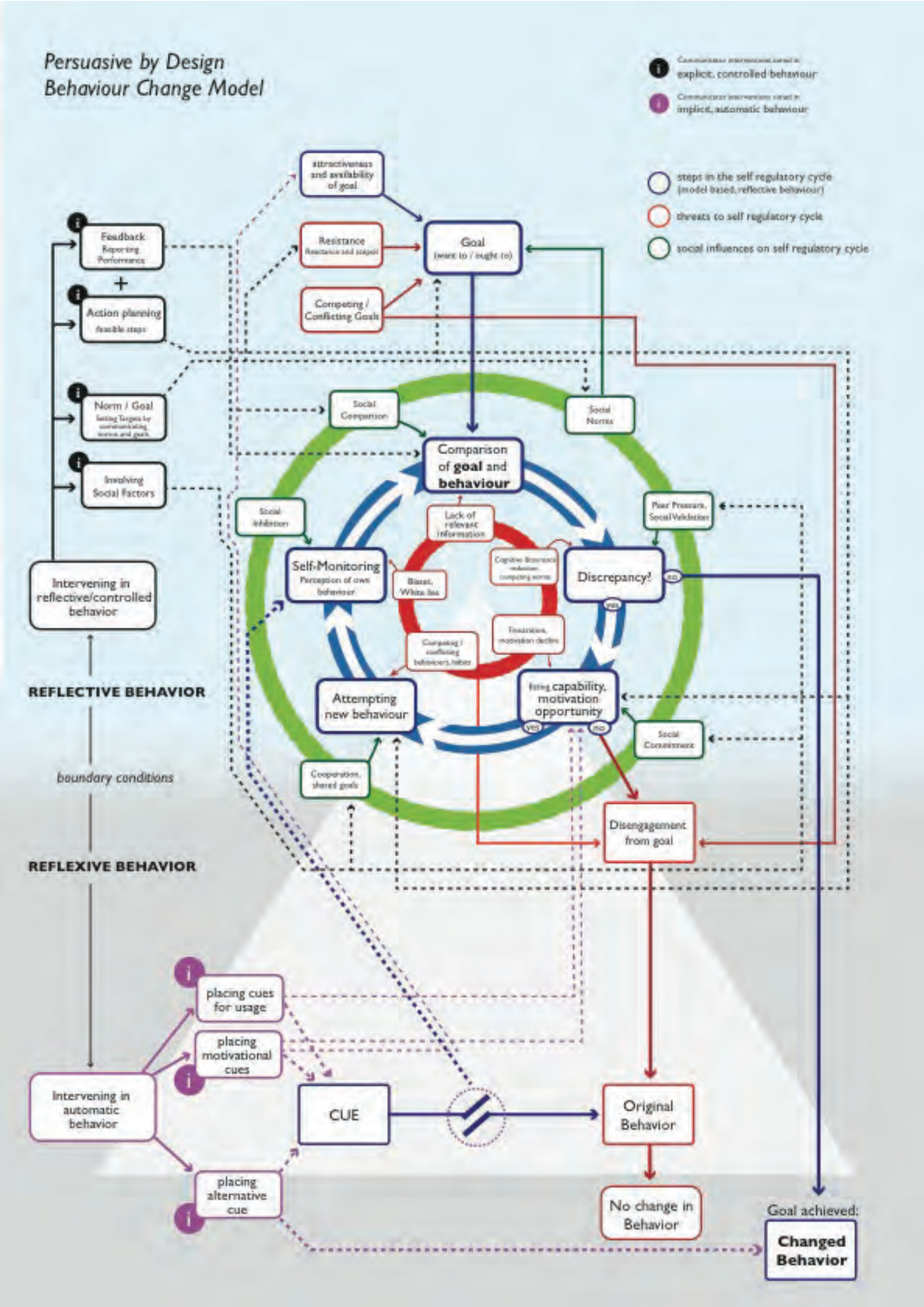

Fig. 1: The Persuasive by Design-Model

For this reason, we developed and tested a workshop during three field sessions. Two of these sessions were aimed at developing interventions to reduce car use at rush our times, with a total of eight participants working on four different concepts. Another session, with twelve participants, was held at a large energy-distribution company, where an intervention is being developed to reduce the CO2-footprint of their fleet of service vans.

Questionnaire results, interviews and participatory observation during the workshops revealed that participants were enthusiastic about the model and the insights provided. Participants were confident about its usefulness both in developing the current concepts the workshop helped developing and in future work. Both questionnaire results and observations showed that after the introduction, participants were able to better identify strengths and weaknesses of their concepts and improve their concepts accordingly.

Research plan for further development and testing

In the months after this conference, we will prototype and test a further interactive version of the tool. This version will enable participants to use the model without having to participate in a time-consuming workshop setting. This tool will be tested in further co-design sessions with design professions in which we provide the tool and ask them to design an intervention to address a real-world problem. We will employ participatory observation and qualitative research methods to evaluate the utility of the tool in these sessions. Finally, we will test the model in an experimental setting, comparing the tool's effectiveness with other existing models and tools.

REFERENCES

1. Zachrisson, J., & Boks, C. (2012). Exploring behavioral psychology to support design for sustainable behavior research. *Journal of Design Research*, *10*(1), 50-66.
2. Michie, S., Ashford, S., Sniehotta, F. F., Dombrowski, S. U., Bishop, A., & French, D. P. (2011). A refined taxonomy of behavior change techniques to help people change their physical activity and healthy eating behaviors: the CALO-RE taxonomy. *Psychology & Health*, *26*(11), 1479-1498.
3. Jylhä, A., Nurmi, P., Sirén, M., Hemminki, S., & Jacucci, G. (2013, September). Matkahupi: a persuasive mobile application for sustainable mobility. In *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication* (pp. 227-230). ACM.
4. Michie, S., Abraham, C., Whittington, C., McAteer, J., & Gupta, S. (2009). Effective techniques in healthy eating and physical activity interventions: a meta-regression. *Health Psychology*, *28*(6), 690.
5. Van Woerkum, M.A., Aarts, N., (2012). *Accountability: New challenges, new forms.* Journal of Organisational Transformation & Social Change, Volume 9, Number 3, November 2012 , pp. 271-283(13)
6. Pettersen, I., & Boks, C. (2008). User-centred design strategies for sustainable patterns of consumption. In T. Geer Ken (Ed.), *Proceedings: Sustainable Consumption and Production: Framework for action, Brussels, Belgium. Conference of the Sustainable Consumption Research Exchange.*
7. Ajzen, I. (2011). Theory of planned behavior. *Handb Theor Soc Psychol Vol One*, *1*, 438.
8. Rosenstock, I. M. (1974). The health belief model and preventive health behavior. *Health Education & Behavior*, *2*(4), 354-386.
9. Fogg, B. J. (2009, April). A behavior model for persuasive design. In *Proceedings of the 4th international Conference on Persuasive Technology* (p. 40). ACM.
10. Dolan, R. J., & Dayan, P. (2013). Goals and Habits in the Brain. *Neuron*, *80*(2), 312-325.
11. Verplanken, B., & Wood, W. (2006). Interventions to break and create consumer habits. *Journal of Public Policy & Marketing*, *25*(1), 90-103.
12. Verbeek, P. P. (2006). Persuasive Technology and Moral Responsibility Toward an ethical framework for persuasive technologies. *Persuasive*, *6*, 1-15.
13. Gardner, B., Whittington, C., McAteer, J., Eccles, M. P., & Michie, S. (2010). Using theory to synthesise evidence from behaviour change interventions: the example of audit and feedback. *Social science & medicine*, *70*(10), 1618-1625.
14. Lewin, K. (1939). Field theory and experiment in social psychology: Concepts and methods. *American journal of sociology*, 868-896.

chi
sparks
2014

Pointing Task Evaluation of Leap Motion Controller in 3D Virtual Environment

Joanna C. Coelho
LIACS, Leiden University,
Niels Bohrweg 1, 2333CA Leiden,
The Netherlands
joannacamargo@gmail.com

Fons J. Verbeek
LIACS, Leiden University,
Niels Bohrweg 1, 2333CA Leiden,
The Netherlands
f.j.verbeek@liacs.leidenuniv.nl

ABSTRACT

Performing tasks in virtual environments are to increasing extent becoming normal practice; such is possible due to the developments in graphic rendering systems and interaction techniques. Application areas from entertainment to medical industry benefit from gestural 3D interaction. With this in mind, we set out a study aiming to research the relevance of using determined 6DoF input devices in interacting with three-dimensional models in graphical interfaces. In this paper we present an evaluation of 3D pointing tasks using Leap Motion sensor to support 3D object manipulation. Three controlled experiments were performed in the study, exposing test subjects to pointing task evaluations and object deformation, measuring the time taken to perform mesh extrusion and object translation. Qualitative data was gathered using the System Usability Scale questionnaire. The data show a strong correlation between input device and performance time suggesting a dominance of the Leap Motion gestural interface over mouse interaction concerning single target three-dimensional pointing tasks. Multi-target tasks were performed better with mouse interaction due to issues of 3D input system accuracy. Performance time regarding shape deformation task demonstrated that mouse interaction outperformed 3D Input device.

Author Keywords

3D object manipulation; pointing task evaluation; 3D input device; leap motion.

ACM Classification Keywords

HCI, gesture interaction, 3D environment, user evaluation

INTRODUCTION

Despite of developments in 3D graphics rendering systems, we still face a lack of knowledge when it comes to interaction with three-dimensional environments. For the construction of this type of interaction it is important to consider a system that permits the user to manipulate the 3D objects in the most natural possible manner, as naturalness directly influences the system usability along with the engagement that the user might present during interaction [18].

Three-dimensional virtual objects and environments can be controlled in various manners, for example by making use of 2D or 3D input devices, providing the user with three, six or more degrees of freedom for translating and rotating objects [11]. Nowadays, the most common scenario for 3D virtual object manipulation are 3D graphics rendering systems on simple desktop setups; this makes the interaction possible, however not yet optimal [14]. More sophisticated Virtual Reality systems tend to use 6DoF sensors, which can be described as 3D input devices that enable translation and rotation (pitch, yaw and roll) in all three axes (cf. Figure 1). Such devices are used to measure position and orientation of limbs providing three-dimensional data regarding the user's movement.

Figure1. Graphical description of movements addressed in 6DoF input devices

Even though we have observed a lot of technical development in the interaction field within the last two decades, 6DoF interaction is still challenging due to limitations of sensor technologies, not enough knowledge on how humans interact with computer generated 3D environments and the recurrent task-specific demands and constraints of each interaction device itself [7]. A few of the most well-known areas that benefit from three-dimensional interaction on virtual environments are: 3D modeling and scene composition, visual programming, medical visualization, prototyping, designing for engineering purposes, browsing large datasets, Technology Enhanced Learning and real-time 3D communication such as the Web3D [7]. Researching the aspects of 3D interaction is relevant as there are many appliances that might vary for every field. Nevertheless, the general goal of the casual user is usually related to browsing, manipulating or interacting with three-dimensional data. With the latter in mind, the purpose of this study is to compare user performance between interacting with a mouse and a Leap Motion

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

device; subsequently, to find out whether or not the use of the Leap Motion device is beneficial for manipulation of virtual objects.

RELATED WORK

Although much work has been done in the field of 3D object manipulation, there are several aspects of interaction that still need to be further analyzed. Early research in the field was conducted with the aim of evaluating 3D input devices in the context of 3D interaction techniques and its relation to user performance [23]. Due to the constant development of interaction devices and rendering systems such research is still common practice.

To evaluate 3D input devices acting in virtual environments, researchers often use Fitts's Law to understand and predict user's reaction time in relation to pointing tasks. A study by Kouroupetroglou, G. et al. [10] reports on a pointing task evaluation comparing between mouse and Wii Remote Control input devices. The study was divided in 2D and 3D experiments in which both Wii Remote and mouse conditions were tested. The two-dimensional experiments were performed in a plane virtual environment counting with 16 circular targets equidistantly arranged from the starting point while, in the 3D case, 8 spherical targets were positioned on the vertices of a cube. The results gathered from both conditions showed that the Wii Remote was outperformed by the mouse in 2D and 3D pointing tasks. It is important to note, however, that the response of the Wii Remote, and thereby its interaction, was reported troublesome with certain light conditions.

Another study by Raynal et al. [16] defends the importance of unifying 3D pointing task evaluation, based on the ergonomic requirements stated in the ISO 9241-9 standard. In this study, researchers adapt the standard evaluation protocol of input devices for 2D pointing tasks, considering important variations that a 3D environment might imply. The devices used for the experiment are the 3D mouse *Space Navigator* and the *Polhemus Patriot* motion tracking input system. One of the most striking adaptations concerns to the validation of reached target in the context of pointing task. It is stated in the ISO 9241-9 that the validation is successful once the cursor is within the target's width. These authors proposed, however, that collision with the target already entails the validation of target reached. This results in a much more positive index of performance by the users and reinforces the necessity of occasional adjustments in pointing task evaluation on 3D-environments.

In [2], Bérard et al. conducted two experiments aiming to investigate the dominance of the mouse in desktop 3D interaction in relation to other 3D input devices. In this research, the *mouse*, *DepthSlider*, *SpaceNavigator* and *Wii Remote* are used as input devices. Evaluation was accomplished by measuring user-performance time when completing pointing tasks inside of a virtual cubic environment. In addition, in the attempt to analyze the bio-signals of the participants, researchers recorded data of galvanic skin response (GSR), heart rate (HR) and volume pulse amplitude (BVP). The experiment demonstrated that the mouse was more efficient than the other devices for accurate placement. In this research is was also concluded that the more degrees of freedom, the worse the performance time for task completion while the stress measured on the user tends to be higher. Nonetheless, it still remains unclear whether the interaction design of the experiment negatively influenced the results of the research in terms of 6DoF input devices.

NATURAL USER INTERFACES

Gestural interfaces are based on recognition and mathematical interpretation of gestures performed by the user, resulting in interactive scenarios that vary in relation to case-specific tasks depending on the goal of the interaction designer. Such interfaces are part of a group of input systems denominated Natural User Interfaces, or NUI. Natural User Interfaces can be classified in two main groups that can be ergonomically distinguished in relation to the physical contact with the body of the user: *wearable* and *touchless* interfaces. As the name suggests, wearable interfaces can be defined as input devices worn by users that contain sensors or markers in order to capture motion with the desirable precision. Systems such as the *Dataglove*, *MOVE* and *WiiMote* can be considered wearable Natural User Interfaces. Touchless interfaces, on the other hand, are characterized by the lack of physical contact with the human body, enabling the user to draw commands without having to touch any equipment. Devices under this category can be essential for determined 3D tasks such as sterile image-guided surgery, once again reinforcing the importance of researching the usability of such devices. Working examples of touchless NUI are the *Microsoft Kinect*, *ASUS Xtion Pro Live*, and the *Leap Motion* sensor.

In this research, we have chosen to use the *Leap Motion* sensor (cf. Figure 2) to perform the experiments with our test subjects. The Leap Motion device combines infrared LEDs and two cameras under a black glass, enabling the software to track finger movements as they are moved over the sensor. The decision to test this device in detriment of others was determined by its commercially announced qualities such as portability, purported accuracy and ease of use, suggesting its possible popularization in the context of domestic 3D environments and virtual object manipulation setups.

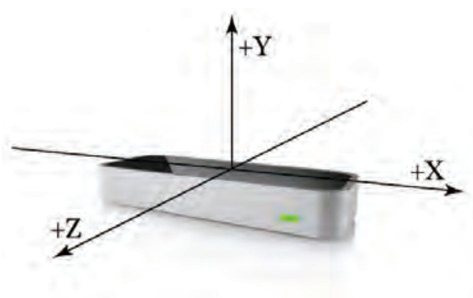

Figure 2. The Leap Motion sensor

3D OBJECT MANIPULATION

A few authors provide us with surveys and comparisons of distinct interaction techniques, describing the main functions that these input devices perform in their respective virtual environments. Chris Hand [7] reports three main operations that the application fields, which profit from 3D interaction, usually make use of, namely: *object manipulation, viewpoint manipulation and application control.* In this paper, we will focus on object manipulation, keeping in mind that, in future work, the other two main tasks should be investigated. According to Subramanian [18], the essential atomic actions within object manipulation can be described as selection, translation and deforming. In this study we will focus mainly on translation and deforming aspects, as we will further illustrate in our experiments.

Our aim is to draw conclusions about the system performance through measurements made during user interaction and therefore it is important to elucidate what variables to take into account for the analysis of the executed tasks. In his study about user performance in relation to input devices, Zhai [23] defines six usability aspects for a 6DoF input device; i.e., *speed*, *accuracy*, *ease of learning*, *fatigue*, *coordination* and *device persistence*. Among all these characteristics of three-dimensional input interaction we will quantitatively measure speed and ease of learning, while accuracy coordination and device persistence are known variables inherent to the given system. Fatigue will be measured qualitatively with the help of the System Usability Scale (SUS) [3].

METHOD

The variety of 3D input and interaction techniques has resulted in many different methods that are utilized to evaluate the performance of the user. Consequently, the novel characteristics of specific input devices might require the creation of ad-hoc approaches for 6DoF interaction evaluation techniques. From the literature, two major approaches are commonly observed as related to the field: structured approach and ad-hoc approaches. In brief, we can describe the structured approach as a composite of methods that aim to assess the pointing task data in a structured manner usually based on Fitts's model. The ad-hoc approaches may vary for case-specific tasks and devices. In this paper we preferred to make use of the structured approach along with inferential and descriptive data analysis in order to evaluate Leap Motion and Mouse input devices in relation to the proposed experiments. Qualitative measurements were performed using the System Usability Scale (SUS), which was filled in by the test subjects right after completion of all tasks.

Experiment Design

Test subjects were randomly divided into two groups that were exposed to different conditions related to the type of input device. The control group was exposed to the mouse condition while the experiment group performed its tasks with Leap Motion gestural interface. An optical mouse of 5V and 100mA wired to the computer via USB was used in the experiment. The sensitivity of the mouse was kept consistent during the whole experiment, not being specifically adjusted for every subject independently.

Subjects from control and experiment groups were exposed to the same virtual environment and target positions only differ on their input interaction method. Reaction time was measured in all the tasks and the initial pointing position of the user as well as target coordinates are known and equal for all test subjects. The experiment task environment was programmed in *Processing.js* supplemented by the *Onformative Library* in order to enable the gestural interaction. Overall, 35 subjects were tested, being 20 in the experimental group and 15 in the control group.

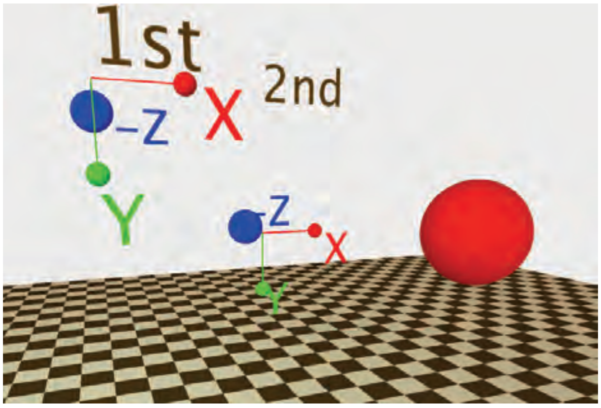

Figure 3. Experiment 1

In order to observe the correlation between input device and effectiveness of object translation we designed two 3D pointing tasks that were performed by the users in a given three-dimensional virtual environment (cf. Figure 3). It is important to notice, that in both cases, viewpoint or camera manipulation was not enabled, providing the test subject with a single angle of vision in order to make decisions with respect to their spatial movements. This decision was taken with the aim to isolate the distance and time variables, keeping in mind that *viewpoint manipulation* should be explored in future work. In the first pointing task, test subjects were instructed to reach a point in space by positioning a red colored sphere onto the first denominated target. The target was described to the user as the "intersection of all axis". The second task had two targets demanding the user to position the sphere on target 1 and subsequently on target 2. Please note that trigonometric and statistical analysis regarding the second pointing task is calculated considering the trajectory from point one to point two and not from the starting point. The validation of target selection is defined within a field of 60 voxels and once the sphere in positioned partially or completely within the field, a console message returns the time taken to reach the target, in milliseconds.

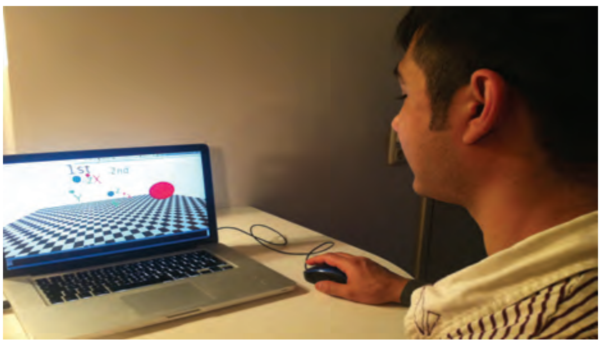

Figure 4. Mouse Interaction

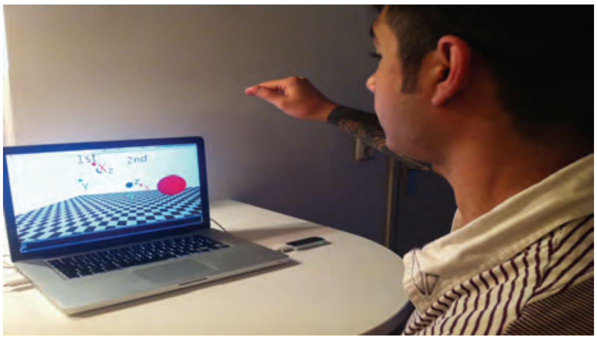

Figure 5. Leap Motion Interaction

Both pointing tasks were analyzed according to variations of Fitts' Law in order to measure the *Index of Performance* (cf. Equation 1) of the given tasks in relation to their *Index of Difficulty* (cf. Equation 2).

$$IP = ID/MT \quad (1)$$

$$ID = \log2(D/W+1.0) \quad (2)$$

$$MT= a+b*\log2(D/W+1.0) \quad (3)$$

The first equation can be described as a formula used to calculate the *Index of Performance* or throughput of a pointing task. Equation 2 aims to calculate the index of difficulty of each pointing task where *D* is the distance between starting point to the center of the target and *W* is the width of the target. Since in our experiment all given targets had the same dimensions, the distinction between the two different given indexes of difficulty was determined by target distances. Equation 3 can be used to predict time measurements concerning the pointing task, where *a* and *b* represent empirical constants determined through linear regression

Although the conventional Fitts's Law is commonly used in research and multidimensional design tasks, the calculation applies only to one-dimensional movements, compromising the comprehension of three-dimensional data, when it comes to manipulation of virtual objects in 3D environments. Due to our different starting point, we adapted Fitt's law for work in a 3D environment, where *c* is an arbitrary constant to be determined through linear regression and θ is the angle between the starting point and target according to Figure 6. One can see in Equation 5 the adaptation we made considering the terms abovementioned in Equation 2. Therefore,

$$ID_3=\log_2(s/b+ 1.0)+c \sin \theta \quad (4)$$

thus changed to:

$$ID_3=\log_2(D/W+ 1.0)+c \sin \theta \quad (5)$$

Considering the referred adaptation of the Fitts's Law to three-dimensional tasks, indexes of difficulty were calculated considering several values of *c* as indicated in [15] and a constant target width.

After completing both calculations, we could understand that the three-dimensional version of Fitts's Law explains more clearly why the second target is harder to reach than the first one since this formula takes into account the angle expressed in a two-dimensional frontal plane from starting point in relation to the target, differentiating indexes of difficulty not only considering distance and width of target but also the referred angle.

Arbitrary constant	TASK 1			TASK 2		
c values	θ	D	ID	θ	D	ID
0	135°	19	3,45	315°	12	5,35
0.1	135°	19	3,52	315°	12	5,28
.0.2	135°	19	3,59	315°	12	5,21
0.3	135°	19	3,66	315°	12	5,14
0.4	135°	19	3,73	315°	12	5,07
0.5	135°	19	3,8	315°	12	5
0.6	135°	19	3,87	315°	12	4,93
0.7	135°	19	3,94	315°	12	4,86
0.8	135°	19	4,01	315°	12	4,79
0.9	135°	19	4,08	315°	12	4,72
1	135°	19	4,15	315°	12	4,65

Table 1. Variation of *c* value on adaptation of Fitts's Law for analysis of three-dimensional tasks

In addition to the first two pointing tasks, a third task concerning 3D object modeling was developed with the aim to evaluate the overall performance of the subjects from the two input conditions while deforming a 3D shape (cf. Figure 6), therefore calculating the average reaction time in both situations. This task consists of re-shaping a deformed cube by extruding one face of the object. The interaction was designed by moving the cursor or tracked hand on a specific axis.

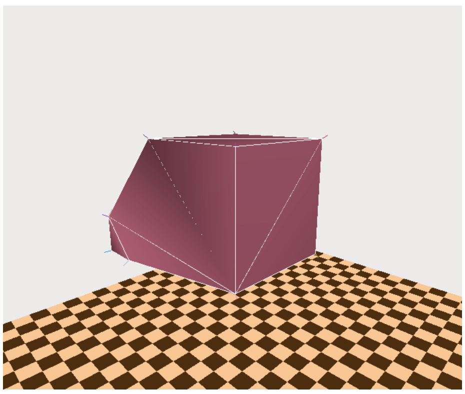

Figure 6. Experiment 2

Experiment Procedure

The mouse condition group was submitted to a brief instructional video, since a pilot study revealed that a few users did not comprehend how to perform the tasks. The 30 seconds of audiovisual demonstration were followed by the completion of the three tasks and data logging of the tasks.

Subjects in the experimental group were also exposed to a short video containing instructions on how to perform the pointing tasks. However, unlike test subjects exposed to mouse condition, the experimental group underwent a short training period that was performed individually. Each subject was introduced to the Leap Motion gestural interface by performing two minutes of interaction with a 3D environment specifically designed for learning purposes. In this environment, users did not have a pre-determined task, therefore interacting freely with a wired white sphere, being able to control the 3D position and rotation of the given shape. After getting acquainted with the gestural interface interaction in the context of a 3D virtual world, subjects were asked to perform two experiments concerning pointing tasks and object modeling experiment.

RESULTS

We tested with 15 participants for the mouse condition and 20 participants for the gestural interface condition. Between the 35 participants, we have 23 male and 12 female subjects from different ages and nationalities (cf. Figure 7). In Figure 8 you can observe the distribution of participants under both conditions by age groups.

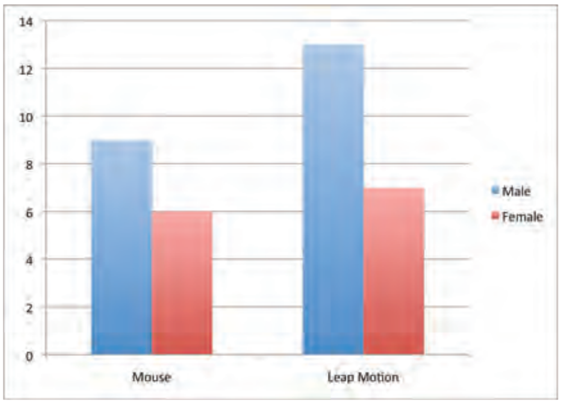

Figure 7. Number of participants in mouse and Leap Motion conditions by gender.

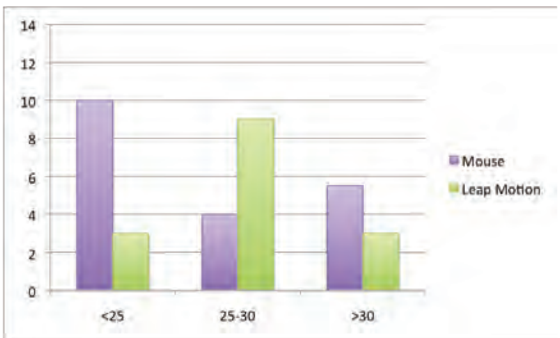

Figure 8. Number of participants in mouse and the Leap Motion conditions per age group.

The Mouse Condition

In Figure 9 the learning rate of the mouse condition is illustrated, showing that, at a second attempt, the user takes approximately 3 seconds less than the first time to reach the same target. The moderated learning curve is expected since we assumed that the mouse input interaction is (already) mastered by all users.

The Leap Motion Condition

The learning curve concerning the Leap Motion device is more accentuated since this is a practically unknown input device within our sample population. However, the interaction device is quite user-friendly, enabling performance time differences of even 8 seconds less in the second attempt to reach target than in the first.

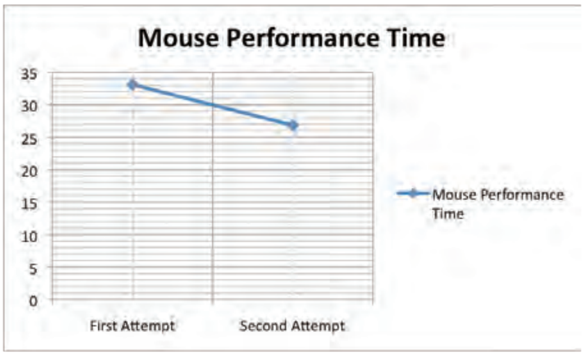

Figure 9. Learning progress on mouse condition

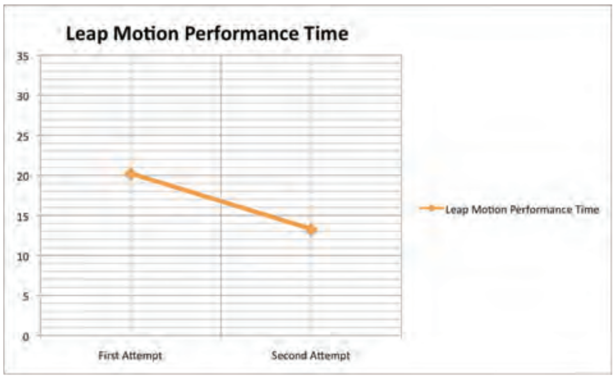

Figure 10. Learning progress on Leap Motion condition

Comparison between Conditions

To assure significance of the given values, a measure independent *t-test* was performed in both conditions for all three given tasks. In the first pointing task, we found significance in the performance time scores for Leap Motion (M=20.25, SD=8.96) and mouse (M=33.09, SD=16.99) conditions; $t(30)=2.41$, $p = 0.05$. The second pointing task showed the following t-test results regarding Leap Motion (M=4.88, SD=2.6) and mouse (M=2.46, SD=1.32) conditions; $t(30)=3.59$, $p = 0.05$. The third task, which involved mesh extrusion, did not achieve the minimum rate required by the t-test due to its high standard deviation: Leap Motion (M=19.45, SD=73.77) and mouse (M=33.09, SD=16.99); $t(30)=0.4$, $p = 0.05$.

Task 1

The first pointing task performed in this experiment has only one target describing a movement from starting point to target, meaning that there are no obstacles or other tasks within this trajectory. As we might observe, under the given constraints, 3D input interaction outperforms mouse interaction regarding the first pointing task (only one target).

After analyzing performance time means, correlation was found between gender and task completion time, showing that females outperformed males in the first pointing task in both conditions (cf. Figure 12).

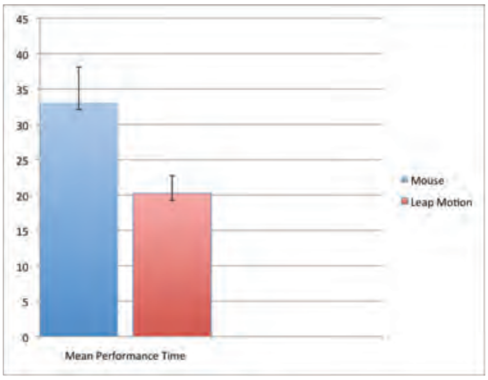

Figure 11. Comparing overall performance time means in both conditions (Task 1)

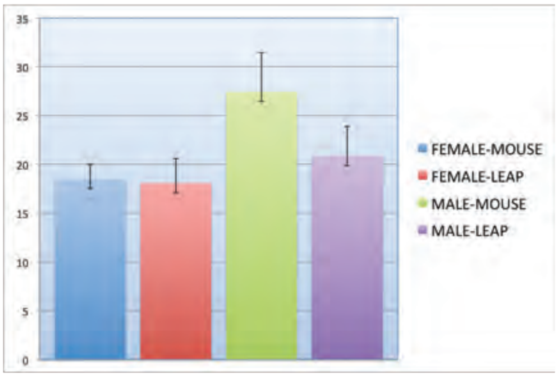

Figure 12. Comparison between performance time means in Task 1 for both conditions by gender

Below we can see a comparison between Leap Motion's and Mouse's learning curves concerning a single target task and considering performance time means regarding first and second attempt to reach the same target. Observing Figure 12 we can conclude that the mean performance time of the Leap Motion device shows much faster performance time compared to the mouse condition.

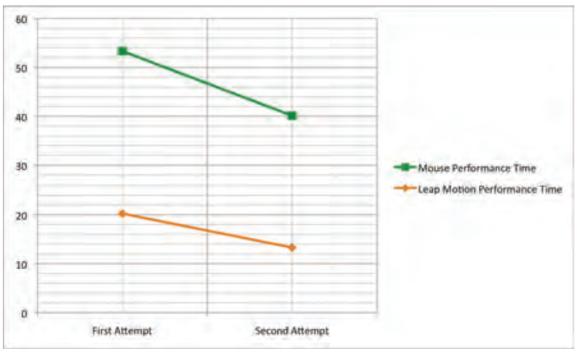

Figure 13. Comparing learning process of both devices

Task 2

The second pointing task contains two targets, assuming a trajectory described along starting point, 1st target and 2nd target. The value we considered in the data analysis and geometric calculations is equal to the spatial difference between target 2 and target 1. Unlike the first pointing task, Task 2 showed faster performance times in the mouse condition. In our case this might suggest that the additional degrees of freedom inherent to the gestural interface might be misleading when consecutively aiming at targets with different "z" coordinates rather than aiming at one single target. It is important to remember that viewpoint manipulation was disabled and that in mouse condition, the "z" axis could be assessed through the roll of the mouse while in the Leap Motion condition the third dimension is achieved by finger-tracking.

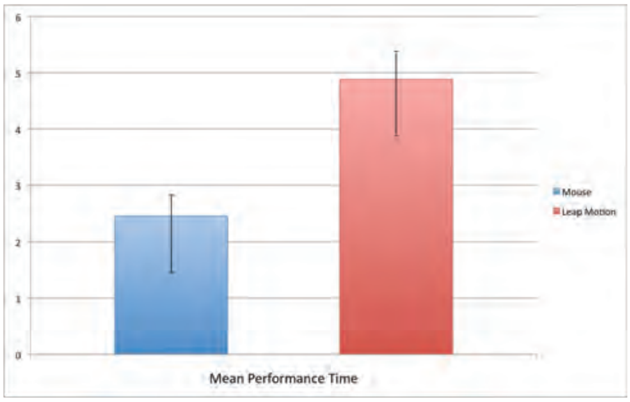

Figure 14. Comparing overall performance time means in both conditions (Task 2)

As we might observe, male and female subjects had a similar performance time in both conditions of Task 2. No significant difference was found between performance time and gender distinction.

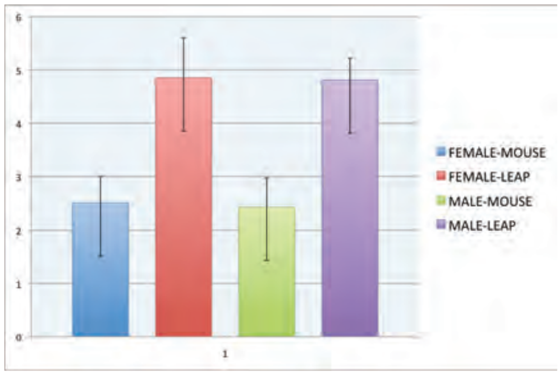

Figure 15. Comparison between performance time means in Task 2 for both conditions by gender

Comparison between qualitative measurement scores and performance time showed correlation between shorter task completion times and higher scores on the System Usability Scale, in which the Leap Motion condition scored higher, indicating a better satisfaction with the device. We must, however, consider that there might be a novelty effect caused by the unfamiliarity of the subject with the device, making subjects score higher on qualitative evaluations due to their interest in such new technology.

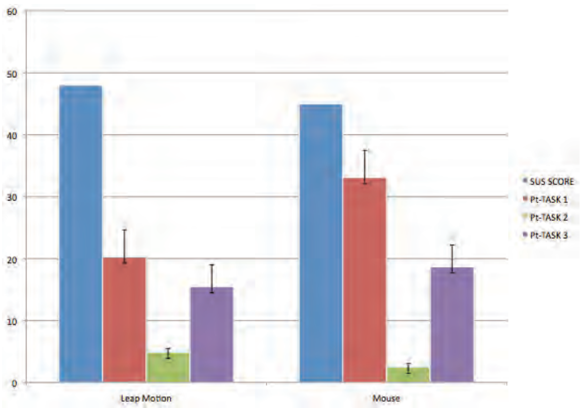

Figure 16. Relation between performance time means for each task and device compared to SUS score means

DISCUSSION

This study shows a comparative performance evaluation of pointing task interaction, showing that although 3D input interaction is qualitatively very well rated by the participants, accuracy is still an important issue that in more dynamic virtual environments can greatly compromise performance time.

From our sample population we could conclude that single target tasks were really simple to perform in the virtual environment with the gestural interface, but the same did not happen once multiple targets were arranged in the experiment. This can be explained by the fact that the second target was located behind the first target and the z-axis was accessed through mouse-roll interaction on the mouse condition, which is still much more accurate than the tracking performed by the Leap Motion, showing that inaccuracy of the tracking can dramatically compromise performance time.

Interestingly, gender correlations were found showing that the females outperformed males in the first pointing task regarding performance time.

CONCLUSIONS AND FUTURE WORK

Based on the results from the experiments, we can conclude that, within the constraints of the tasks developed in this research, the presented 3D input device outperformed mouse interaction only in single target situations, showing that 3D translation is less cumbersome when the “z” axis is provided as input based on real-life movement mappings. However, accuracy issues can prejudice performance time of more complex spatial movements (multiple targets).

Negative aspects of using the 3D input device for complex spatial interactions were reported in the development stage. Device accuracy issues are one of the biggest challenges for the popularization of those 3D input devices. Still concerning 3D tasks, expert users have shown in both quantitative and qualitative studies to be extremely biased towards mouse interaction, electing the mouse as the most reliable and practical device for manipulating 3D objects.

Further investigations and experimentations into *viewpoint manipulation* and *application control* is strongly recommended, since that would provide us with more clear guidelines on how to fully interact with a given 3D software by means of 3D input devices, including window and menu navigation, state changes and camera control.

In order to assess the performance of other available 3D input devices when modeling and manipulation 3D virtual objects, further research should be guided considering a broader selection of 6DoF input systems, enabling a more complete overview of the advantages of one technique over of a second, or third one. It is interesting to point out that making an assessment of the weak aspects from the evaluated 3D input systems could contribute with the development of existing or novel interaction devices.

ACKNOWLEDGMENT

Our deepest appreciation goes to the test subjects whose time and effort were innumerably valuable throughout the course of this study.

REFERENCES

1. Balakrishnan, R., Baudel, V, Kurtenbach, G., Fitzmaurice, G. The rockin’mouse: integral 3d manipulation on a plane. *CHI ’97* (1997).
2. Bérard, F., Benovoy, J.Ip.M., El-Shimy, D., Blum, J.R., and Jeremy R. Did "Minority Report" get it wrong? Superiority of the mouse over 3D input devices in a 3D placement task, *INTERACT (2), Springer,* (2009), 400-414.
3. Brooke, J. SUS: A Quick and Dirty Usability Scale. *Usability Evaluation in Industry*, (1996).
4. Chen, M., Mountford, S. J., Sellen A. A study in interactive 3-d rotation using 2-d control devices. *SIGGRAPH ’88: Proceedings of the 15th annual conference on Computer graphics and interactive techniques*, (1998), 121–129.
5. Froehlich, B., Hochstrate, J., Skuk, V., Huckauf, A. The globefish and the globe-mouse: two new six degree of freedom input devices for graphics applications, *ACM Conference on Human Factors in Computing Systems (CHI), ACM Press* (2006), 191–199.
6. Haan, G. Techniques and Architectures for 3D Interaction. TU Delft, (2009).
7. Hand, C. A survey of 3D interaction techniques, *Computer Graphics Forum 16*, (1997), 269-281.
8. Hinckley, K., Tullio, J., Pausch, R., Proffitt, D., Kassell. N. Usability analysis of 3d rotation techniques. *UIST ’97: Proceedings of the 10th annual ACM symposium on User interface software and technology*, (1997), 1–10.
9. Jacob,R. J. K., Sibert, L. E., McFarlane, D. C., Preston ,J. M. Integrality and separability of input devices. *ACM Trans. Comput.-Hum. Interact., 1(1),* (1994), 3–26.

10. Kouroupetroglou, G., Pino, A., Balmpakakis, A., Chalastanis, D., Golematis V., Ioannou N., Koutsoumpas, I. Using Wiimote for 2D and 3D pointing tasks: gesture performance evaluation. *Gesture Workshop , Springer*, (2011), 13-23.

11. Kulik, A., Hochstrate, J., Kunert, A., Froehlich, B. The influence of input device characteristics on spatial perception in desktop-based 3D applications. *3DUI , IEEE*, (2009), 59-66.

12. Kunert, A., Huckauf, A., Froehlich B. A comparison of tracking- and controller-based input for complex bimanual interaction in virtual environments. *B. Froehlich, R. Blach, and R. van Liere, editors, EG IPT-EGVE 2007*,(2007), 4352.

13. Lee, J., Boulanger, C.N. Direct, spatial, and dexterous interaction with see-through 3D desktop. *SIGGRAPH Posters , ACM,* (2012), 69.

14. Liang, J., Green, M. JDCAD: A Highly interactive 3D modeling system. *3rd International Conference on CAD and Computer Graphics, Beijing, China*, (1993), 217-222.

15. Murata, A., Iwase, H. Extending Fitts' Law to a three-dimensional pointing task, Human Movement Science, (2001) 20 791–805.

16. Raynal, M., Dubois, E., Schmitt, B. Towards unification for pointing task evaluation in 3D desktop virtual environment. South CHI 2013: 562-580.

17. Silveira, W. G. Manipulation of 3D objects in collaborative environments Using the Kinect Device. *Federal University of Uberlândia*. (2009).

18. Subramanian, S., Ijsselsteijn, W. Survey and classification of spatial object manipulation techniques. *IPO, Center for User-System Interaction, Eindhoven University of Technology*, (2000).

19. Ware, C. Using hand position for virtual object placement. *The Visual Computer 6(5),* (1990), 245–253.

20. Wuthrich, C.A. An analysis and a model of 3D interaction methods and devices for virtual reality. *Proceedings of the Eurographics Workshop*, (1999),18-29.

21. Zhai, S. Human performance in six degrees of freedom input control. *Ph.D. Thesis. Univ. of Toronto*, (1995).

22. Zhai, S. Interaction in 3D Graphics. *SIGGRAPH Computer Graphics Newsletter 32,* (1998).

23. Zhai S. *User performance in relation to 3D input device design*. ACM Computer Graphics 32(4), (1998), 50–54.

How Turn-Taking Influences the Perception of a Suspect in Police Interviews

Rifca Peters
University of Twente
Enschede, The Netherlands
post@rifca.nl

Merijn Bruijnes
University of Twente
Enschede, The Netherlands
m.bruijnes@utwente.nl

Rieks op den Akker
University of Twente
Enschede, The Netherlands
h.j.a.opdenakker@utwente.nl

ABSTRACT

We study turn-taking behaviour in non-cooperative dialogue for the development of believable characters in a serious game for conversational skill learning in the police interview context. We describe a perception study to see how participants perceive a suspect's interpersonal stance, rapport, face, and deception when the turn-taking of the subject varies. We influence the perception of the suspect's stance by altering the timing of the start of speech with respect to the ending of the interlocutor's speech. The results of the study contribute to the development of an embodied conversational agent capable of natural human-system conversation with appropriate turn-taking behaviour.

Author Keywords

Embodied Conversational Agent; Turn-Taking; Serious game; Social skill training; Police interview; Believable virtual humans; Experimental perception study

ACM Classification Keywords

I.2.7 Natural Language Processing

INTRODUCTION

In human conversation we try to adhere to a "one-at-a-time" approach. Sacks, Schegloff and Jefferson [15] proposed a systematic, offering a set of rules to provide next-turn allocation to one interlocutor and thereby minimizing gap and overlap. However, moments of overlapping speech or silences occur frequently in human conversation [16]. These silences and moments of overlapping speech are often communicative in their own right [5,11,14]. Emotions and the stance people take towards each other influence turn-taking behaviour. "A clash of opinions also means a clash of turn-taking" [12]. Contrary to the dynamic turn-taking behaviour in human conversation, turn-taking behaviour in current natural dialogue systems is often restricted by a "one-at-a-time'' rule. Conversational agents (CAs) are limited to listening or speaking and listening is initiated either on a place predetermined by the system or whenever the user makes a sound, resulting in an unnatural human-system interaction. Exceptions are the dialogue systems that allow more free turn-taking behaviour [18].

In the context of the COMMIT P2 project we are working towards a computational model for human-like suspect turn-taking behaviour. This model supports the creation of a believable embodied conversational agent (ECA). This ECA will be used in a social skill training serious game for police officers that is currently under development. Rich CA turn-taking behaviour, including pauses, interrupts, and hesitation, is expected to support a more natural human-system interaction. A previous conversation analysis [2] showed that a factor such as the topic of conversation influences the interpersonal stance and the turn-taking behaviour of the suspect. Moreover, the stance of the suspect appeared to be related to the interpretation of suspect silences, e.g., a silent response from a suspect with a positive stance is interpreted as timidity while during a hostile stance it is related to withdrawal. Turn-taking strategies seem to have an effect on the perception of the agent [12,13].

The purpose of this study is to investigate how turn-taking behaviour influences the impression that observers get from a suspect in simulated police interviews. We look at the relation between turn-taking behaviour and perception of power, affiliation, rapport, face, and deception. We use extracts of police interviews in which we systematically vary turn-taking behaviour to study the influence of turn-taking on perception. This study focuses on the police interview setting. Police officers receive training on recognition and strategic use of interactional phenomena such as dominance [3]. Due to this experience, their perception of affective stance may be different from untrained people. The results of this work will inform the creation of a serious game that *police officers* will use to train their interview skills.

This paper is organized as follows. In the next section, we give a brief overview of relevant literature. Next, we present the Research Question and describe Methodology of the perception study. We conclude with a Discussion of the expectations of the results and the relevance of the results for the development of a conversational agent.

RELATED WORK

Literature on theoretical frameworks of and results from conversation analysis on turn-taking in police interviews provided us with some suggestions on which factors influence turn-taking behaviour in police interviews.

Yoong [24] showed that police officers interrupt suspects to prevent them from turn completion. These deliberate interruptions are considered signs of assertion of power

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

[14,24]. Due to the asymmetric question/answer adjacency pairing, a police interview is structured to provide the officer with control over the conversation [4]. Haworth [7] claimed that power is under constant negotiation and reported recognition interrupts, minimal responses, taking extended turns, and interruptions of question as techniques used by suspects to access control in police interviews. Vrij [22] suggest that truth tellers adopt a "tell all" approach resulting in a talkative mood opposed to liars who adopt the "keep it simple" approach resulting in a less talkative mood. A more in-depth analysis of silence during stages of deception and truthfulness is given by Komter [10] who suggests that resistance by evasion or defence is a sign of deception and silences after a statement or question are associated with a non-contradicting position of the suspect. This absence of denial is often highlighted by an officer by allowing a long silence. To be considered relevant, denial should be provided immediately following or interrupt an accusation [9]. Rapport is considered a critical step in eliciting trust and building a relationship in professional interaction and therefore a prerequisite for techniques used in police interviews, e.g., to get cooperation from the interviewee [1,23]. Suspects tend to talk more openly in harmonious interactions and cooperation and agreement are increased. Discomfort –considered a lack of rapport– is displayed by stretches, fillers and pauses in the speech of the suspect [6]. In turn-taking we adhere to the terminology proposed by Heldner and Edlund [8], distinguishing two silences: gap and pause, two overlaps: between and within speaker, and bridged turn transitions: a smooth transition with no discernable silence (less than 0.18s) (see Figure 1). The type of question can influence the perception of an utterance. For example, a question directly addressing the suspect requires a response while this is not necessary for a statement. Also, an open-ended question is expected to be followed by an extensive response while yes or no are satisfactory responses for a closed question [17]. The type of question asked is related to the function of a question, e.g., information seeking for open-ended questions and conformation seeking for closed questions [20]. Moreover, case-related question may be more sensitive than small talk.

Ter Maat et al. [12,13] show that the manipulation of turn-taking strategies can lead to different perceptions of an agent on personality scales, interpersonal scales, and emotional scales. They conclude that these strategies can be used in the repertoire of expressive behaviours of agents reflecting these dimensions. We extend on this perception study. Based on the literature review, we hypothesize there is a relation between turn-taking behaviour and perception of power, affiliation, rapport, face, and deception.

RESEARCH QUESTION

To support the development of a computational model for turn-taking behaviour of a virtual suspect agent we evaluate the suggestions presented in the literature review: we assess if turn-taking behaviour is indeed related to the perception of interpersonal stance and investigate possible interaction between factors of interpersonal stance. The main research question is: '*What influence do variations in turn-taking behaviour have on the perception of power, affiliation, rapport, face and deception of a virtual suspect?'*. We formulated hypotheses following the same pattern for each of these factors: a turn-taking feature influences the factor. For deception (the other factors are omitted to conserve space):

In interactions with audible pause between sequential suspect turns, the suspect is perceived as more deceptive than in latched sequential turns.

In interactions with a gap between a question from an officer and the answer by a suspect, the suspect is perceived as more deceptive than in latched or overlapping question/answer adjacency pairs.

In interaction with a gap between a statement by an officer and a denial by a suspect, the suspect is perceived as more deceptive than in latched or overlapping denial.

METHODOLOGY

We selected extracts from our police interview corpus [2] and generated them with variance in timing of the start of speech with respect to the ending of the speech of the other interlocutor. These extracts are presented to participants who are asked to fill in a short survey on their perception on the personality, emotional state and interpersonal stance of the suspect after each extract. A pilot study is conducted to evaluate the stimuli and survey.

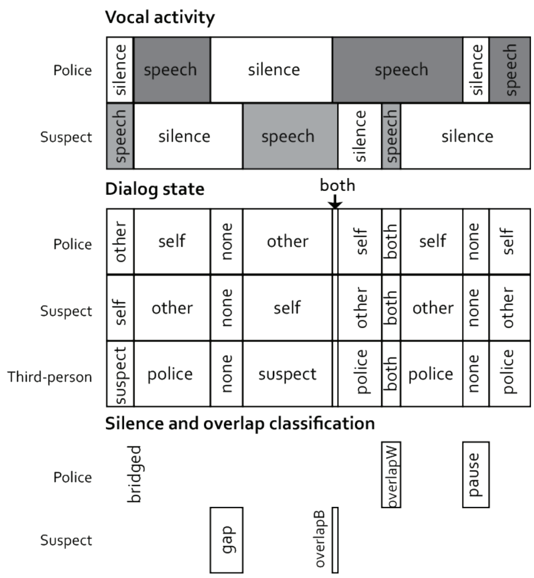

Figure 1: Top: Vocal activity of two speakers. Middle: The dialogue state shows who is speaking (depending on the perspective). Bottom: Classifications of the dialogue state: gap, pause, between-speaker overlap, within-speaker overlap, and bridged speaker transitions.

Participants

Police officers or police trainees are our participants (n=30) as their perception of affective stance may be different from untrained people due to their experience. Participants need to be native or proficient Dutch language users as all stimuli are in Dutch.

Stimuli

The stimuli, extracts from the corpus of Dutch police interview training videos [2], are generated using Ivona (ivona.com) text to speech. To maximize recognition of both speakers they are of opposing gender and the gender of the officer and the suspect are counterbalanced over all stimuli. All stimuli are generated using a single male and a single female voice. The extracts selected from the corpus demonstrate –or contradict– one interpersonal factor, see Figure 2. For each extract an altered version is created in which the turn-taking behaviour is adjusted while maintaining the content of the conversation as much as possible. Names are replaced by fictive names of similar length. Utterances are recorded and edited to vary the turn-taking using Audacity (audacity.sourceforge.net).

Design and Procedure

Participants are seated in front of a computer with loudspeakers. On the computer an online survey is presented. The participant is provided with information about the study and ensured confidentiality of their data. On each page the participant plays an audio file. Each file consists of an extract of a simulated conversation between an officer and a suspect. To distinguish between the officer and the suspect both interlocutors are of opposing gender.

To gather how a suspect is perceived, a survey is presented after each stimulus. The survey is the same for each stimulus –except for gender that is altered to comply with the gender of the speakers in the extract under assessment– and consists of opposing statements pairs to be rated on a 7 point semantic difference scale for: *dominance*, *friendliness*, *togetherness*, *cooperativeness*, *positivity*, *agreeability*, *attentiveness*, *politeness*, *respectfulness, autonomy, closeness, resistance, compliance* and *deceptiveness*. The chosen scales include the characteristics of interpersonal stance [3] and the factors of rapport [19]. Questions are counterbalanced for polarity where possible.

DISCUSSION

Previous research investigating police interviews included some aspects of silence or interruption and provided us with

```
410  P  till what time?=
412  S  =till eeeh twelve o'clock I had lessons
        (0.7)
414  P  hmm hmm
        (0.4)
416  S  ·Hhhh, then, eh, I went into the city for
        a bit with a classmate
        (.)
418  P  ok
        (.)
420  S  because eh
        (0.52)
422     yeah we also eh kind of eh needed things
        for in hair, I also do hairstyling, so we
        also needed things to put in hair and then
        I went home
```

Figure 2: Example of the transctiption (translated from Dutch) of one stimulus for deception demonstrating the "tell all" approach. The officer asks a question (line 410) a response is immidiatly provided by the suspect (line 412), the suspect volunteers extended information and repeatedly self-selects (lines 416, 420-422).

suggestions on how personality, emotional state, and interpersonal stance influence turn-taking behaviour in a police interview setting [1,4,6,7,9,10,14,22,23,24]. However, these studies included turn-taking as one aspect within overall suspect behaviour and where not directed at the development of a model for turn-taking behaviour of a suspect. In this study we investigate if the factors influencing turn-taking according to the literature hold for a suspect in Dutch police interviews. We investigate whether variations in turn-taking behaviour lead to differences is the perception of the suspect. The first results will be presented at the Chi Sparks conference.

We expect the results of the study to contribute to the understanding of underlying factors influencing the (unconscious) choices a suspect makes if and when to speak. This understanding of underlying factors is needed to create an embodied conversational agent capable of mimicking human-like turn-taking behaviour which will support a more natural conversation between a human and an ECA. It can show its internal state by showing the appropriate turn-taking behaviour. For example, a virtual suspect in a dominant stance will display behaviour such as interrupts or when the agent has a deceptive stance it will take shorter turns and longer pauses in storytelling. See for an example of this type of agent [21]. The current study will try to determine what appropriate turn-taking behaviour is given the internal state of the agent that it tries to convey.

A potential limitation of our study is the usage of auditory-only stimuli. This removes the interference of non-verbal behaviour. However, non-verbal behaviour is undoubtedly important for the perception of an agent and will be available in the intended game environment. Studies on the perception of ECAs that incorporate verbal and non-verbal agent behaviour are required in the police domain. Also, all stimuli are short extracts (between 25 and 40 sec). However, longer extracts might be necessary for observers to form a consistent perception of the speakers. To the best of our knowledge, no research is done to investigate the relation between vocal stimuli length and perception agreement. In [2] we saw that inter-annotator agreement was low for short fragments, but we showed that global patterns become evident over longer periods.

By investigating the influence of turn-taking behaviour on the perception of a virtual suspect in police interview we aim to support the development of a virtual suspect for use in a social skill training serious game for police officers. By assessing the influence of turn-taking behaviour on the perception we gather knowledge about the extent of importance to model turn-taking behaviour and the appropriate behaviour given a desired interpersonal stance.

ACKNOWLEDGMENTS

This research is supported by the Dutch national program COMMIT.

REFERENCES

1. Abbe, A. and Brandon, S.E. The Role of Rapport in Investigative Interviewing: A Review. *J. Investig. Psych. Offender Profil. 10*, (2013), 237–249.
2. Akker, R. op den, Bruijnes, M., Peters, R., and Krikke, T. Interpersonal stance in police interviews: content analysis. *Computational Linguistics in the Netherlands 3*, (2013), 193–216.
3. Amelsvoort, A. van, Rispens, I., and Grolman, H. *Handleiding Verhoor*. Stapel & De Koning, 2012.
4. Benneworth, K. Police interviews with suspected paedophiles: a discourse analysis. *Discourse & Society 20, 5* (2009), 555–569.
5. Ephratt, M. The functions of silence. *Journal of Pragmatics 40, 11* (2008), 1909–1938.
6. Fogarty, K., Augoustinos, M., and Kettler, L. Re-thinking rapport through the lens of progressivity in investigative interviews into child sexual abuse. *Discourse Studies 15, 4* (2013), 395–420.
7. Haworth, K. The dynamics of power and resistance in police interview discourse. *Discourse & Society 17, 6* (2006), 739–759.
8. Heldner, M. and Edlund, J. Pauses, gaps and overlaps in conversations. *Journal of Phonetics 38, 4* (2010), 555–568.
9. Heydon, G. Silence: Civil right or social privilege? A discourse analytic response to a legal problem. *Journal of Pragmatics 43, 9* (2011), 2308–2316.
10. Komter, M. The Interactional Dynamics of Eliciting a Confession in a Dutch Police Interrogation. *Research on Language and Social Interaction 36, 4* (2003), 433–470.
11. Kurzon, D. The right of silence: A socio-pragmatic model of interpretation. *Journal of Pragmatics 23,* (1995), 55–69.
12. Maat, M. ter and Heylen, D. Turn management or impression management? *Intelligent Virtual Agents,* (2009), 467–473.
13. Maat, M. Ter, Truong, K., and Heylen, D. How turn-taking strategies influence users' impressions of an agent. *Intelligent Virtual Agents,* (2010), 441–453.
14. Momeni, N. Police Genre: Interruption and its Classification as a Sign of Asymmetry in Police Interview/Interrogation. *International Journal of Criminology and Sociological Theory 4, 1* (2011), 615–625.
15. Sacks, H., Schegloff, E. a., and Jefferson, G. A simplest systematics for the organization of turn-taking for conversation. *Language,* (1974), 696–734.
16. Schegloff, E.A. Overlapping talk and the organization of turn-taking for conversation. *Language in Society 29, 01* (2000), 1–63.
17. Snook, B., Luther, K., Quinlan, H., and Milne, R. Let 'em Talk!: A Field Study of Police Questioning Practices of Suspects and Accused Persons. *Criminal Justice and Behavior 39, 10* (2012), 1328–1339.
18. Thórisson, K.R., Gislason, O., Jonsdottir, G.R., and Thorisson, H.T. A Multiparty Multimodal Architecture for Realtime Turntaking. *Intelligent Virtual Agents,* (2010), 350–356.
19. Tickle-Degnen, L. and Rosenthal, R. The Nature of Rapport and Its Nonverbal Correlates. *Psychological Inquiry: An International Journal for the Advancement of Psychological Theory 1, 4* (1990), 285–293.
20. Tracy, K. and Robles, J. Questions, questioning, and institutional practices: an introduction. *Discourse Studies 11, 2* (2009), 131–152.
21. Traum, D. Non-Cooperative and Deceptive Virtual Agents. *IEEE Intelligent Systems,* (2012), 66–68.
22. Vrij, A., Jundi, S., Hope, L., et al. Collective interviewing of suspects. *Journal of Applied Research in Memory and Cognition 1,* (2012), 41–44.
23. Walsh, D. and Bull, R. Examining Rapport in Investigative Interviews with Suspects: Does its Building and Maintenance Work? *Journal of Police and Criminal Psychology 27, 1* (2011), 73–84.
24. Yoong, D. Interactional norms in the Australian police interrogation room. *Discourse & Society 21, 6* (2010), 692–713.

KIDS @ PLAY

chi sparks 2014

Wobble: Supporting Social Play through an Open-ended Play Environment

Alice van Beukering
Department of Industrial Design
Eindhoven University of Technology
a.j.e.v.beukering@student.tue.nl

Linda de Valk
Department of Industrial Design
Eindhoven University of Technology
l.c.t.d.valk@tue.nl

Tilde Bekker
Department of Industrial Design
Eindhoven University of Technology
m.m.bekker@tue.nl

ABSTRACT

In this paper we explore how to design for different levels of social play such as solitary, parallel and group play. For this purpose, we developed an open-ended play environment that supports the three stages of play: invitation, exploration and immersion. This play environment, called Wobble, focuses on socio-dramatic play by triggering children's imagination. Wobble consists of multiple interactive light objects developed for children aged four to six years old. Wobble was evaluated with eighteen children playing in groups of three during a free play session. The results show a clear pattern of how children first approach the play environment in a solitary manner, then explore its interaction possibilities in parallel and eventually become immersed in group play. This pattern can be supported by various design properties as local feedback and spatial interaction rules.

Author Keywords

User-centered design; open-ended play; dramatic play; social play; stages of play; design research

ACM Classification Keywords

H.5.2 [**User Interfaces**]: Interaction styles (e.g., commands, menus, forms direct manipulation), User-centered design.

INTRODUCTION

Within the HCI community, research into play and games is an exciting and evolving field. Research in this field focuses on designing for playful user experiences that often have a societal or personal impact. Especially for children, play is an essential learning activity through which they practice new skills and explore imaginary worlds [1]. Play is generally considered to fulfill an important aspect in the development of children [9, 10, 19]. In play, children can develop an imaginary and temporary world with flexible rules and boundaries [7]. Children's play activity can take on a variety of forms. When involved in games, a structured form of play, rules and goals are predefined. When involved in an unstructured form of play, children have more freedom to create their own play. In our research we want to primarily support the latter form of play. Therefore, we focus on the design of play environments that support next to social and physical play also open-ended play [4]. In open-ended play, there are no predefined rules or goals. Players are allowed to create their own rules, goals and games [18]. A play environment can support open-ended play by offering flexible and dynamic interaction opportunities that are multi-interpretable depending on the imagination of the players [4]. In this way, open-ended play supports the creativity and imagination of the players.

In this paper we focus on dramatic play, which is still a rather unexplored type of play within the field of open-ended play environments. In dramatic play children engage in a pretend activity by dramatizing life situations, taking a role of someone else, or by bringing life to an inanimate object, i.e. making a doll talk [12]. In the ages of 4 to 6 years old children engage in dramatic play, while at the same time their interest towards social play scenarios increases [9]. The Play Observation Scale (POS) [12] identifies three levels of social play: solitary, parallel and group play. Additional research is needed to understand how an open-ended play environment can support these different levels of social play. For instance, by identifying design properties that can help designers create more socially engaging play environments. Therefore, in this paper we look in more detail at different levels of social play within dramatic and open-ended play.

In previous work, we have seen that in open-ended play the experience of interaction goes through three stages: invitation, exploration and immersion [17]. In the *invitation stage* the potential players are attracted to the play design through their senses from a distance. When players start to explore interaction opportunities and try to understand rules and affordances of the play design, players enter the *exploration stage*. Subsequently, in the *immersion stage* players are involved in ongoing engagement of the actual play experience. For example, players show strong expressive behavior while playing or focus on achieving a clear goal in the form of parallel or group play. The relevance of designing for these three stages has been shown [17]. We aim at further strengthening this approach by also taking into account the different types of social play.

We created a new open-ended play environment called Wobble (see Figure 1). Wobble is an interactive play environment, intended for children in the ages of 4 to 6 years old. The environment consists of multiple interactive objects in the shape of balls on a flexible stem. Children can interact with the balls and push them around. The objects sense their movement with an accelerometer and react to

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

this with light feedback (RGB LEDs). The objects are connected with wires to an Arduino that controls the accelerometers and LEDs. Children can interact with each object individually, but the objects can also communicate with each other. Some balls in the system will softly pulsate and lights will jump from one ball to another. The play objects respond with light feedback when pushed, adding a sense of "living creatures" inside the balls. When a child subtly pushes an illuminated ball, it changes color, from turquoise to purple or the other way around. When a child pushes the ball a little harder the light will jump to another randomly assigned ball, as if the "living creatures" fly away. The play objects aim to trigger the children's curiosity and fascination (invitation stage). When exploring the play objects, the children may wonder about what can happen (exploration stage). Children are stimulated to develop their own worlds of fantasy while engaged in dramatic play (immersion stage).

In this paper, we explore the following research question: "*How do different design properties of an open-ended play environment support different levels of social play: solitary, parallel and group play; during the stages of play: invitation, exploration and immersion?*" To be able to examine this research question, we performed an explorative user study with Wobble. In the rest of this paper, we first discuss related work in the area of designing for play. Then, we describe the design process and concept of Wobble and relate this design to levels of social play and stages of play. We continue with a description of the set-up and methodology of our explorative study. After that, we present the results of this study addressing a verification of the concept, levels of social play and stages of play. The paper is concluded with a discussion.

Figure 1: The Wobble prototype.

RELATED WORK

The research question for our study resulted from a literature study in play design research. Play has been described based on many different dimensions, including social, emotional, motor and cognitive dimensions, but also based on play contexts and structural properties of play behavior [8]. For the research described in this paper, work on socio-dramatic or pretend play, and on social play is most relevant. Below, we describe an illustrative, non-exhaustive overview of previous designs for play, and where relevant the related studies, in the areas of open-ended play, dramatic and social play and stages of play.

Open-ended play

ColorFlare [5] is an example of an open-ended play design (see Figure 2). ColorFlare consists of multiple handheld objects that can be rolled and shaken and reacts to these actions with different types of light feedback. Colors can be transmitted between objects to stimulate communication between players. The multiple interaction opportunities of ColorFlare can lead to the creation of a diverse variety of games [5]. This shows potential for designing for open-ended play, and we aim to elaborate on this research with a focus on both social and dramatic play.

Dramatic play and social play

Children at the ages of 4 to 6 years old are mainly interested in fantasy and pretend play. In this type of play "magic" often plays an important role. "Magic" can serve as an explanation for things that children are not yet able to understand [1]. Herein, "magic" can be the bridge between the play world of children and their cognitive development. Thus, the component of magic to support socio-dramatic play of children playing together is an important property to embed in an open-ended play environment for children.

There are a lot of commercial available products on the market that support dramatic play in children, i.e. baby dolls, such as the *BabyBorn* [3], kitchen and store toys, dress-up items and accessories and many more. However, these examples do not integrate technology. The *ActiMates™ Barney™* [2, 15] is an example of a play design that integrates technological advantages and focuses on dramatic play. The design is an animated plush doll in the shape of a familiar media character. Based upon how children bring life into inanimate objects mimicking social interactions, Barney fosters dramatic play by invoking such responses in children. Barney is an example of a quite concrete and structured play design that shows that dramatic play can be supported through integrating technology in a play design. Another example that focuses on dramatic play is the *StoryMat* [13] (see Figure 2). The design focuses on collaborative storytelling by offering a play space that is able to record and recall children's own narrating voices and the movements of the children's toys on a mat.

Figure 2: Two examples of related work: ColorFlare [5] and StoryMat [13].

Concerning social play, two theories have inspired our work: those by Parten [9] and Broadhead [6]. Parten [9] defined the degree of play participation in six sequential social participation categories: unoccupied behavior, solitary play, onlooker behavior, parallel play, associative play, and cooperative play. The theory developed by Broadhead describes the various social play behaviors in more detail. She created a methodology called the Social Play Continuum [6], in which social play behavior is measured by the level of reciprocity in language and action. The

Continuum describes four social domains: associative play, social play, highly social play, and cooperative play. These theories provide starting points for examining variations in play behavior ranging from more individual to more social behavior, and examining behavior in terms of children's actions and language.

An example of a design for social play is *FeetUp* [11], which encourages the practice of social skills. The design is a playful accessory embedded in the children's shoes. FeetUp provides audiovisual feedback when the child jumps or is off the ground. When multiple children have the same accessory, they have a common interest and can develop shared goals. When children share goals they have to explain ideas, argue, and negotiate, through which they practice social skills. These examples show evidence of how interactive play designs can support dramatic and social play in children. But these examples are not designed with open-ended play in mind, except for FeetUp, or combine their focus on social and dramatic play. The play designs that do focus on social play do not focus on specific levels of social play. Besides, the play designs focusing on dramatic play are still rather concrete and structured. In our study we will examine how to support dramatic play through a more abstract intelligent play design.

Stages of play

To date we are not aware of any other play design besides our own work that is designed explicitly with the three stages of play in mind. In a previous study we designed the *FlowSteps* [17] that focuses on the three stages of play. The design consists of multiple, flexible mats that provide light feedback when pressure is applied to a mat. It attracts players by randomly lighting up one of the mats (invitation stage). Players individually or together can explore possible interaction opportunities offered by the mats (exploration stage). Subsequently, players are able to create their own rules, goals and games by giving meaning to the output modalities of the mats (immersion stage). An explorative study conducted with the FlowSteps has generated insights on how to design for playful experiences during the three stages of play. The results show that the design properties that support playful experiences can vary for the different stages of play, e.g. some experiences are more important when players enter the immersion stage. In this paper we build further upon this work [17], by examining how an open-ended play environment supports different levels of social play during the three stages of play.

DESIGN RESEARCH APPROACH

We followed a research through design approach [20]. According to this approach, knowledge is generated by creating and testing prototypes throughout various iterations. Theoretical assumptions support the creation of a design, which is evaluated in context to verify underlying assumptions. Both aspects of design assumptions and user experiences are crucial in gathering qualitative and situational insights. The levels of social play, the three stages of play and open-ended play form the theoretical background and inspiration for the unique design of Wobble.

Design Process

Wobble was developed in three iterations. Each iteration consisted of developing or improving the prototype, evaluating it with children from the target group (see Figure 3) and reflecting on this. In the first iteration, the prototype (see Figure 3: left) consisted of three interactive flexible balls on a stick providing simple light feedback. The prototype was successful in evoking diverse interactions and open-ended play scenarios. However, an early evaluation showed that the prototype was still rather abstract for children from the target group (4-6 years old) and a more specific trigger for the development of dramatic play is needed. In the second design iteration, a new prototype (see Figure 3: right) with a larger amount of interactive balls was developed with light feedback as well as background sounds. This prototype offered a more specific context through its physical design, i.e. little felt insects were attached to the balls. A second evaluation showed that this prototype was successful in evoking diverse forms of interaction and play scenarios and in supporting dramatic play. For example, some children believed that the balls could be controlled by creating "wind" either by blowing or by waving with their hands. Furthermore, we observed that children played in different social settings with Wobble. For instance, children played solitary with one play object, or together with multiple objects, i.e. children pushed the lights from ball-to-ball together. These pilot observations showed that children engaged in different levels of social play. In a third design iteration the final prototype was developed, which consists of multiple interactive light objects. The final prototype was improved on sensor accuracy, physical construction, i.e. by aggravating the foot of the objects to increase stability, and fine-tuning the interaction rules, i.e. changing light color when subtly tapping the ball. Figure 4 shows an illustration of the final interaction rules.

Figure 3: Children playing with the first prototype (left) and the second prototype (right) of Wobble.

Design Rationale

Different design properties of Wobble support various levels of social play. Wobble consists of multiple independent objects that children can play with individually. In **solitary play** the child "engages in an activity entirely alone, usually three feet away from other children" [12]. In the design of Wobble, a child can play with one object in the system. The independent objects offer local interaction rules, i.e. subtly tapping the ball to change its light color. Children engage in **parallel play** when they "engage in activity beside, but not with other children, usually at a distance of three feet or less" [12]. The objects are grouped together at a distance of approximately three feet. This allows children to play in parallel with one independent object while imitating each other's explored interactions. Different objects in the system provide different states of light and color feedback, i.e. always two objects are on, while 3 objects are off. The child can notice these differences, which can trigger the child to become somewhat attentive to his or her playmates and/or to start to engage in parallel speech ("the child verbalizes his or her own thoughts for the benefit of other children" [12]). In this way, the design incorporates properties that can direct children to parallel play situations. The child engages in **group play**, if he or she "engages in an activity with another child or children, in which cognitive goal or purpose is shared among all members" [12]. The interaction rules between the objects in the system can direct children to group play, as it invites children to play together with the multiple objects. For example, the children can push one illuminated ball to make the light jump to another randomly assigned ball in the environment.

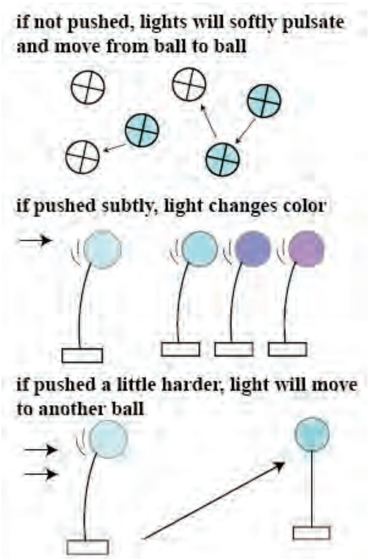

Figure 4: Final interaction behavior of Wobble.

Expectations

Based on pilot observations, we expect that the three levels of social play (solitary, parallel and group play) will be supported during the three stages of play in a pattern as illustrated in Table 1. When engaged in play with Wobble, we expect children to move from solitary, to parallel, to group play as children move from the invitation, to the exploration, to the immersion stage. Children will mostly approach the design in a solitary manner, when persuaded to play with the design in the invitation stage. In the exploration stage, children will move towards parallel play situations. When children enter the immersion stage, we expect children to become more engaged in group play, i.e. children will come up with shared games and goals in a group. This does not mean that other combinations of social play and stages of play will not occur, but we expect them to be less likely to occur. For example, in principle a design could elicit mostly solitary play in all three stages, because it provides less opportunity for children playing together.

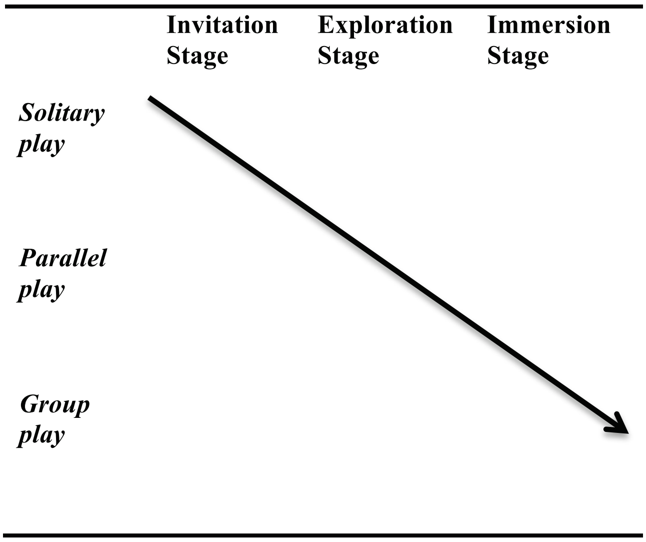

Table 1: Expectations about how levels of social play will be supported during the three stages of play.

STUDY

We performed an explorative study as a first examination of how to support levels of social play within open-ended play. The set up of this study is rather small-scale, with a focus on collecting qualitative, detailed and rich information on how children play with Wobble. In this section, we describe the research question, set-up, methodology and analysis of our study.

Research Question

We aim to explore the overall research question: "*How do different design properties of Wobble support different levels of social play: solitary, parallel and group play; during the stages of play: invitation, exploration and immersion?*". To illustrate the play quality of Wobble, we also observe how children play with the design.

Set-Up, Methodology and Analysis

The study was conducted at a primary school in The Netherlands with eighteen children, eight girls and ten boys, aged 4 to 6 years old. Six groups of three children, some same-gender and others mixed gender, played in a free play session of 15-20 minutes with the design. Wobble was placed in a familiar environment: an unused classroom at the school. The children were divided in groups by their teacher based upon the likeliness of how well the children play together. All sessions were carried out according to the same protocol. A session started with guiding the children to the classroom where the design was set up. After entering the classroom, the test leader left the children alone with the design for about half a minute, to evaluate the *invitation stage*. Next, the test leader gave the children a short introduction and invited the children to explore the design. After five minutes, the interaction was further explained and the children were asked to come up with a game. When the children got distracted or started talking to the test leader their attention was brought back to the design.

The play sessions were analyzed using a qualitative research method [14]. During each session video recordings were made. All play sessions were coded for relevant events. In a first round of analyzing the videos, the first author gained experience using the POS [10] to code levels of social play. In this analysis, we did not use the POS coding scheme in a quantitative manner (i.e. counting for how long children were involved in a type of social play). Instead, we observed in a qualitative manner how children played in a session and coded that with a certain type of social play. In a second round of observation, the first author coded the levels of social play in the three stages of play. A second independent researcher, after becoming acquainted with the POS [12], coded one video on levels of social play and stages of play. Both researchers discussed their findings until consensus was reached. In a third round of observation attention was paid to the specific design properties that support different levels of social play. The second independent researcher acted as a discussion partner for analyzing how different design properties support different levels of social play.

RESULTS

In this section, we present the findings of our study. First, we describe the game play during the free play sessions with Wobble. Then, we describe how particular design properties supported different levels of social play in the three stages of play. Note that all quotes from the children were translated from Dutch to English.

Game Play

Upon seeing the design, the children reacted enthusiastically and were curious to start playing with the design. Some children were a bit more hesitant, but the active lights of the design easily persuaded them to begin playing. During the free play sessions the children engaged in various forms of play. All children actively explored the design trying out various interaction possibilities such as pushing, clapping, waving or blowing to control the lights inside the balls. A large amount of children became engaged in dramatic play, e.g. one child made graceful 'wizard-like' movements trying to control the lights, saying: "*Pouf, pouf, pouf, turn on!*" Besides this, children came up with their own games, e.g. one group of children played a game like "*musical chairs*", in which each child tried to catch an illuminated ball. Most games involved rules about trying to turn the lights in the balls on or off. To a lesser degree, the children involved rules about different colors of light in their game play. Furthermore, the game play often involved a spatial element, wherein children went from object-to-object, e.g. children engaged in a game of pushing the blue light(s) from ball-to-ball. Children also played with the design without using the interactivity of the design, i.e. children explored the physical properties of the objects. Moreover, children were involved in conversations, communicating their actions and observations with each other, e.g. "*Look, if you tickle the ball it will wake up!*"

Although the game play differed between children, most children engaged in fairly similar play dynamics. At first children's game play was quite calm, but over time it became more active, e.g. running from object-to-object.

Figure 5: Children playing with Wobble.

Overall, Wobble supported diverse types of games engaging a variety of interaction possibilities. On top of the interaction rules of the design, the children created their own meanings, games, goals and rules in multiple ways. In this way, Wobble supported dramatic play and different open-ended play scenarios. Children seemed to enjoy playing with Wobble, expressing playful sounds during their play and explicitly stating they liked playing with the design.

Stages of Play and Levels of Social Play

An overview of findings on the levels of social play and stages of play is shown in Table 2. In the cells of this overview we describe the design properties that support different levels of social play (vertical) in different stages of play (horizontal). The three larger cells with bold text describe the main level of social play in that specific stage of play. The two large arrows show how children primarily move between these different levels of social play and stages of play. The smaller cells with grey text describe how other levels of social play are also supported in that specific stage of play. For example, in the exploration stage the main level of social play is parallel play. But solitary play and group play are also observed, though to a lesser degree. The small circular arrows indicate how children switch between specific levels of social play and specific stages of play. Below, we discuss our findings for each stage in more detail.

Invitation Stage

The video data shows that most children spread over the multiple objects in a solitary manner. Each child has a *personal object* when approaching the play design. Some children observed the objects at a very close proximity before daring to subtly push a ball. Another girl hesitantly walked step-by-step towards an object. Though, in some groups we observed a difference in this pattern. For example, in one group, two children initially tended to move together in the direction of one object in the system. However, when getting close to the object the children chose to move to different objects in the system. Local feedback communicates to the children that the design offers local interaction possibilities for play. Therefore, *local feedback*, i.e. the lights, of the design supported the children's solitary play in the invitation stage. Besides this, solitary play is also supported by the *unfamiliarity of the design*. For example, we observed that most children seemed to need some time to process the play design, expressed in their calm behavior upon seeing the design. Though we observed some exceptions; for example, one boy who enthusiastically started to run in a playful manner between the objects upon seeing the design. Besides this, we observed that most children centered their attention on their own activity of approaching an object, i.e. children did not show true signs of being attentive to other children. Furthermore, in approaching the play design the children did not involve in conversation, and only rarely made statements such as: "*Wow, what are these objects?*"

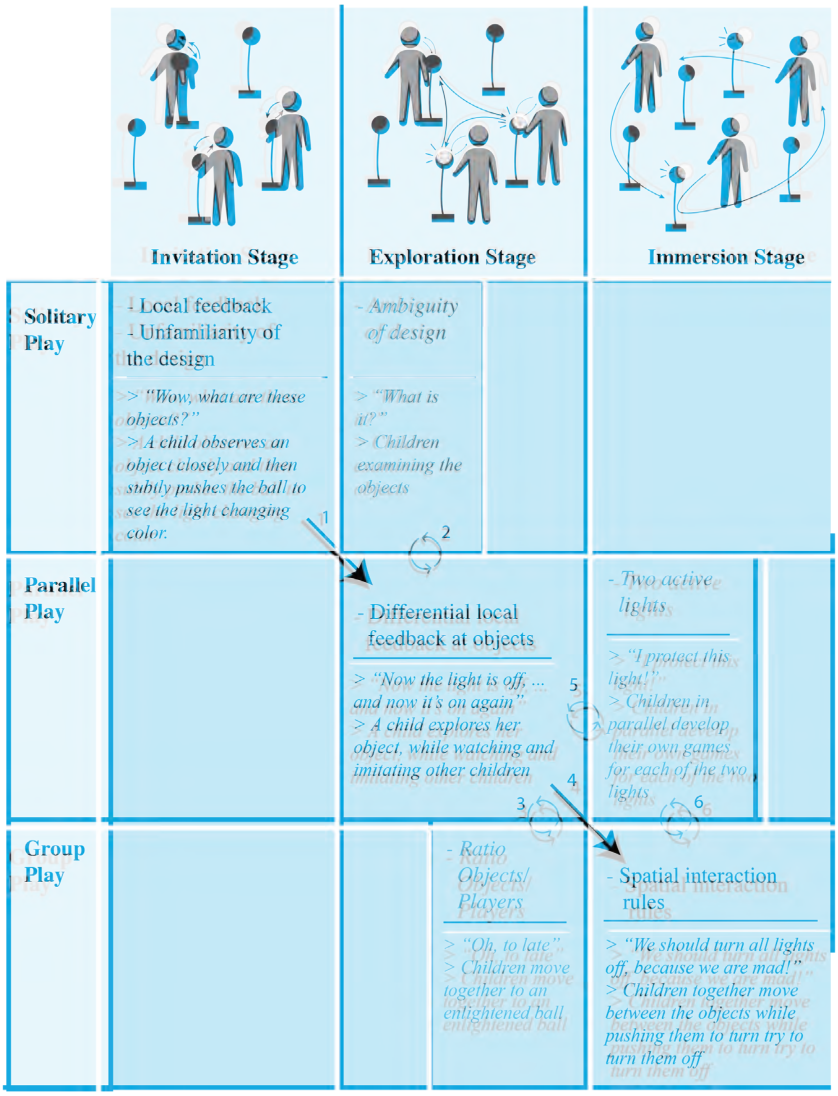

Table 2: Overview of findings: design properties supporting different levels of social play (vertical axis) in the stages of play (horizontal axis). The bold arrows (1 and 4) illustrate how children primarily move between these different levels of social play and stages of play. The small circular arrows (2, 3, 5 and 6) show how children can move between other levels of social play and stages of play.

Overall the play behavior of the children in the invitation stage is solitary. This is supported by the unfamiliarity and local feedback of the design, as well as the fact that each child has a personal object to interact with.

Exploration Stage

When children went from the invitation stage to the exploration stage, children mostly went from solitary play towards parallel play (see Table 2 - *arrow 1*). The different states of local feedback (at the same moment different objects displayed different local feedback) supported the children's parallel play in this stage. Due to the *differences in local feedback* of the objects, children's attention is directed towards each other. Children started to compare and noticed that their object could provide different local feedback in terms of light intensity and color in comparison to other objects in the system. We observed that the children almost directly became attentive to each other when entering the exploration stage. For example, a girl explored her personal object, while sporadically watching and imitating how other children interacted with their personal objects. Children also started to engage in parallel speech:

"verbalizing his/her own thoughts for the benefit of the other children" [12]. For instance, children started sharing their observations: "*Now the light is off ... and now it's on again*", "*I tickle the ball*" or, "*Turn on!*" In the exploration stage, children primarily played individually with their personal objects. However, sometimes the children switched from one object to another. Children tried-out a variety of interactions such as pushing, waving, talking or blowing to control the lights inside the balls. Furthermore, the children engaged in parallel, dramatic play. For instance, one boy knocked at a ball while saying: "*Knock, knock, who is there?*" Another girl examined her ball, after which she said: "*Oh I see, it is a big flower!*"

At the start of the exploration stage, children also engaged in solitary play for relative short periods (see Table 2 - *arrow 2*). For example, due to the *ambiguity* of the design a child became so fascinated by his or her personal object that the child just looked at it for seconds to half a minute. The child lost attention for the other children and did not engage in parallel speech. After this, the child again switched to parallel play.

As the exploration stage progressed, we observed that children sometimes engaged in group play for relatively short periods (see Table 2 - *arrow 3*). Group play in this design was supported by the *ratio illuminated objects versus amount of players* of the design. During the free play sessions, there were always two illuminated objects versus three players in the system. We observed that some children moved together towards an illuminated ball and explored its interaction rules, while engaged in communication. Children did have the general tendency to search for a personal object in the system, whereby children went back towards parallel play scenarios. The children rarely interacted together with one object in the exploration stage.

Overall, in the exploration stage children engaged in parallel speech and were attentive to each other, without sharing cognitive goals or rules. This parallel play was supported by different local feedback at the objects.

Immersion Stage

The video data shows that when children went from the exploration stage to the immersion stage, children generally went from parallel play to group play (see Table 2 - *arrow 4*). We observed that the *spatial interaction rules* persuaded the children to engage in spatial play scenarios, e.g. feedback divided over the balls caused children to move between the objects. Moreover, it triggered children to engage in mutual communication, e.g. children discussed rules, goals and games, and started their sentences with: "*We…*" Together, children started to develop rules, goals or games and in this way entered the immersion stage. For example, a group of children came up with a goal to turn all lights off: "*We should turn all lights off, because we are mad! Oh, you naughty balls!*" When engaged in group play, children directed each other by pointing.

Besides this, although less frequent, children entered the immersion stage in a parallel manner (Table 2 - *arrow 5*). This is supported by the *two active lights* of the design that afford to develop rules, goals and games in parallel concerning one of the two lights. For example, one child was involved in a game of protecting his or her own light, while another child, simultaneously, tried to push another light from one object to another object. Overall, children engaged for relative short periods in parallel play in the immersion stage. Often, children either started to engage in exploring the design again moving back to the exploration stage (see Table 2 - *arrow 5*), or started to engage in group play, in which children developed a shared goal or purpose (see Table 2 - *arrow 6*). For example, we observed that if a child in a parallel play setting came up with a game and communicated this game to other children, they tended to join the game, resulting in group play.

In the video data, we did not observe children moving in groups from the immersion stage to the exploration stage. When the children started to engage in group play, the children actually directly came up with a rule, goal or purpose. For example, a short conversation between two children playing with one ball: a girl said: "*Let's tickle the ball!*" upon which a boy responded: "*Yes, we should wake this light up!*"

Overall, in the immersion stage the children primarily engaged in group play. The children engaged in mutual communication and developed shared goals, rules and games supported by the spatial interaction rules of the design. In this way, children often shared a common goal or purpose in the immersion stage.

CONCLUSION

We were interested in exploring how different design properties of Wobble support different levels of social play: solitary, parallel and group play; during the stages of play: invitation, exploration and immersion. By examining how children play with the open-ended design Wobble we gained a better understanding of how to design for different levels of social play in the three stages of play. The results show that children in their interaction with an open-ended play environment generally move from solitary play to parallel play to group play, as children go from the invitation stage to the exploration stage to the immersion stage. Children primarily approach a play environment in a solitary manner, and subsequently explore its interaction possibilities in parallel at their personal object while being attentive to others and engaging in parallel speech. Children become immersed in a play experience while engaging in group play sharing a common purpose or goal.

Furthermore, the results show how the dynamics of the social play behavior of children develop over time and what specific design properties can support these levels of social play. For example, solitary play in the invitation stage is supported by the local feedback of the design, while group play is supported by the spatial interaction rules in the immersion stage. The results provide insights in how to design for variation in social play behavior of children over time and how to possibly persuade children towards a specific level of social play. Moreover, the results show that, in order to guide children's play to the immersion stage, designs should first offer interaction opportunities for solitary and parallel exploration. This enables children to become immersed (involved in an ongoing engagement) in group play with the play environment.

This study was also a first exploration on how to design for open-ended play with a focus on dramatic play. Designing for this perspective taught us to explicitly take into consideration the abstractness of the design and consider using a specific trigger to stimulate the development of dramatic play.

We expect our findings can help other designers and design researchers in enriching their play designs to support different levels of social play. These findings should not be considered as the pattern children *should* follow when interacting with an open-ended play environment. Instead, we present these results as a design and thinking tool that provides insights in how to design for social dynamics in play. We also believe our findings can be adequate for other design scenarios with a focus on encouraging social interaction.

DISCUSSION

This paper presents a first step in exploring how to design for different levels of social play. Our current findings are based upon one case study and therefore must not be seen or used as final or fixed conclusions. More research into this direction is needed to further ground our results. For instance, future user studies could involve different age groups, as social play behavior tends to be quite age specific. The cognitive capabilities of younger children differ quite a lot from older children, which could influence their social play behavior. Besides this, the number of children influences the group dynamics. Different sizes of groups of children playing with the design can provide further insights into children's social play behavior. Moreover, the personality of children is also interesting to investigate further. For instance, Teele's Inventory of Multiple Intelligences [16] could be used as a tool to describe differences in personality characteristics of children.

Concerning the design of Wobble, some improvement points can also be addressed. The current objects are connected with wires. During the free play sessions some children got distracted by the wires for a moment and asked questions about their functionality. This did not disturb the children's play a lot as their attention quickly focused back on the design. But to avoid these issues we aim for a wireless solution in a next design iteration. Furthermore, the richness of the local and spatial interaction rules could be improved, to support more complex interaction.

ACKNOWLEDGEMENTS

We would like to thank the children who participated in our study and their teachers for their cooperation. This research is part of the Creative Industry Scientific Programme (CRISP), which is funded by Dutch government FES funding.

REFERENCES

1. Acuff, D.S. and Reiher, R.H. *What kids buy and why; the psychology of marketing to kids*. Free Press, Detroit, 1997.
2. Alexander, K. and Strommen, E. "Evolution of the talking dinosaur: the (not so) natural history of a new interface for children." In *Proc. CHI 1998*, ACM Press (1998), 7-8.
3. *BabyBorn*. Retrieved in March 2014 from http://www.zapfcreations.com.
4. Bekker, T., Sturm, J. and Eggen, B. "Designing playful interactions for social interaction and physical play." *Personal and Ubiquitous Computing 14, 5* (2010), 385-396.
5. Bekker, T. and Sturm, J. "Stimulating physical and social activity through open-ended play." In *Proc. IDC 2009, ACM Press* (2009), 309-312.
6. Broadhead, P. *Early years play and learning: developing social skills and cooperation*. RoutledgeFalmer, London, 2004.
7. Huizinga, J. Homo *Ludens: a study of the play element in culture*. Beacon Press, Boston, 1955.
8. Johnson, J.E. Play development from ages four to eight. In: Fromberg, D.P. and Bergen, D. (Eds.). *Play from birth to twelve: contexts, perspectives and meanings*. Routledge, New York, 1998.
9. Parten M.B. "Social participation among preschool children." *Journal of Abnormal Psychology 27, 3* (1932) 243-269.
10. Piaget, J. *Play, dreams, and imitation in childhood*. Norton and Company, New York, 1962.
11. Rosales, A., Arroyo, E. and Blat, J. "FeetUp: A playful accessory to practice social skills through free-play experiences." In *Proc. Interact 2011*, Springer (2011), 37-44.
12. Rubin, K.H. *The Play Observation Scale (POS).* Published by the Center for Children, Relationships, and Culture, University of Maryland, 2001. Retrieved in March 2014 from http://www.rubin-lab.umd.edu/CodingSchemes.
13. Ryokai, K. and Cassell. J. "StoryMat: a play space with narrative memories." In *Proc. IUI 1999*, ACM Press (1999), 201.
14. Strauss, A.L. and Corbin, J.M. *Basics of qualitative research: Grounded theory procedures and techniques*. Sage Publications Inc., California, 1990.
15. Strommen, E. "When the interface is a talking dinosaur: learning across media with ActiMates Barney." In *Proc. CHI 1998*, ACM Press (1998), 288-295.
16. Teele, S. *Teele inventory for multiple intelligences*. Sue Teele & Associates, Redlands, California, 1992.
17. Valk, L. de, Rijnbout, P., Bekker, T., Eggen, B., Graaf, M. de and Schouten, B. "Designing for playful experiences in open-ended intelligent play environments." In *Proc. IADIS GET 2012*, (2012), 3-10.

18. Valk, L. de, Bekker, T. and Eggen, B. "Leaving room for improvisation: towards a design approach for open-ended play." In *Proc. IDC 2013*, ACM Press (2013), 92-101.

19. Vygotsky, L.S. "Play and its role in the mental development of the child." *Soviet Psychology 5, 3* (1967), 6-18.

20. Zimmerman, J., Forlizzi, J. and Evenson, S. "Research through design as a method for interaction design research in HCI." In *Proc. CHI 2007*, ACM Press (2007), 493-502.

chi sparks 2014

BioPong: Adaptive Gaming Using Biofeedback

Dirrik H.G. Emmen
Leiden University
Niels Bohrweg 1, 2333 CA, Leiden,
The Netherlands
dirrik+chi@gmail.com

Georgios Lampropoulos
Leiden University
Niels Bohrweg 1, 2333 CA, Leiden,
The Netherlands
g.lampropoulos@umail.leidenuniv.nl

ABSTRACT

The goal of this study was to develop a biofeedback version of the classic video game Pong in which heartbeat and galvanic skin response measurements are used to adapt the game difficulty according to the physiological state of the users. It was hypothesized that the biofeedback version of Pong would improve user experience and performance. Two prototypes were tested on a total of 12 players. User evaluations have been used to measure user experience and scores have been used to measure user performance. The results show that Pong can be made easier or harder according to the physiological state of the player, which improves user experience. User performance did not improve.

Author Keywords

Biofeedback; video game; adaptive gameplay; heart rate; galvanic skin response; user experience; user performance.

ACM Classification Keywords

H5.2. Information interfaces and presentation: User Interfaces - Evaluation/methodology, Input devices and strategies, Prototyping, Interaction styles. Prototyping, Interaction styles.

INTRODUCTION

In 1972 Atari released the first video game mega hit called Pong. Pong is a simple two player game in which two users hit a ball back and forth between each other by controlling a paddle. In the years after the development of Pong, games changed in graphics, storytelling, and gameplay. One thing that remained is the breadth of emotions a game can evoke [7]. Emotions that arise during a game (e.g., frustration, anger, and stress) can result in unconscious physiological changes, for example a higher heart rate.

The goal of the present study is to test whether the physiological state of the players can be used to improve player performance and experience in the classic video game Pong.

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

Player Versus Player

Player performance during a game depends on previous play experience and the level of expertise reached over time. Imagine a Pong game where one player is experienced and the other one is not. The experienced player will not be challenged by the other player and will most probably win the game. In addition, players can reach a point where they do not feel motivated or challenged anymore. Biofeedback can be used to create a more challenging game experience or to provide a similar gaming experience for people with different experience levels [4; 5].

Biofeedback

Players have different emotions (e.g., boredom and frustration). The heart rate and galvanic skin response (GSR; skin conductance), of a player can provide information about the emotional state that a person is in [3; 8]. For example, the higher the heart rate, the more ecstatic a person is, and the higher the GSR, the more stress a person experiences. These two forms of input are relatively easy to use in video games, because the sensors are small and the costs are relatively low.

Several other games used biofeedback. Some games replaced conventional input by biometric input. For example, in the Atari Mindlink controller muscle activity controls the game and in “Relax-to-win” a player can control the speed of a dragon by relaxing [1]. In “Biofeed the zombies” biofeedback is used to adapt the game environment (i.e., the game was was made more or less scary) to the physiological state of the player.

The general aim of this study was to test whether user experience and performance in Pong could be improved by adaptive gameplay using biometrics.

METHOD

Design

Two prototypes of Pong with biofeedback were developed and tested in two game sessions. The goal of the first game session was to get a better understanding of the physiological state of players during a game and to test the user experience. The prototype was improved after the first gaming session and the improved prototype was tested in the second game session. Each game session consisted of three game conditions that made it possible to compare user experience and performance in normal Pong and Pong with biofeedback. The following three conditions were used: (1) a two player classic Pong game, (2) a two player biofeedback Pong game which became harder when heart rate and GSR increased and easier when heart rate and GSR decreased, and (3) a two player biofeedback Pong game which became easier when heart rate and GSR increased and harder when heart rate and GSR decreased. Each game condition lasted 3 minutes and was played once. Heart rate was mapped to the paddle size. The paddle size became longer (easier) or shorter (harder). GSR was mapped to the ball speed in the player side of the screen. Ball speed either increased (harder) or decreased (easier). The physiological state of the users directly changed the gameplay parameters while playing. Heart rate and GSR categories were made that were mapped on predefined paddle sizes and ball speeds. For the first prototype heart rate values were categorized as followed: 55-65, 65-75, 75-90, 90-120, and all above 120. For the second prototype these ranges were changed to: 55-70, 70-85, 85-100, 100-115, and all above 115. In the first prototype the GSR values were categorized as followed: 0-10, 10-20, 20-30, 30-35, 35-40, and all above 40, and for the second prototype: all below 20, 20-25, 25-30, 30-35, 35-40, 40-45, 45-55, and all above 55.

Participants

In the first game session four users (i.e., two pairs) participated. In the second game session eight users (i.e., four pairs) participated. Player pairs remained the same across the three game conditions in a game session. All participants were in their 20s, and followed the master program Media Technology at Leiden University. In the first game session all participants were male. In the second game session there were four males and four females. In this session each pair consisted of one male and one female.

Measurements

In all three conditions of both game sessions GSR and heart rate were measured. Verbal comments and nonverbal reactions of the players during the game sessions, which were observed by the authors of this paper, were used to measure user experience. To measure user performance the final score per game condition was noted. Game sessions were filmed for documentation.

Hardware and Environment

Game sessions were conducted at Leids Institute of Advanced Computer Science. In both game sessions an Arduino was used to save the GSR and heart rate data from the sensors. In the first game session participants used a keyboard to play the game. In the second evaluation a self-developed controller was used to play the game. The controller was designed in such a way that allowed the GSR sensor to be placed in the backside where players would naturally place their fingers while holding it. A pulse sensor in the shape of an ear-clip was used to measure the heart rate. The controller had a button to start a game and consistent with the original Pong controller, the controller had a potentiometer to control the paddle. The controllers that were used in the second game session are presented in Figure 1.

Procedure

Participants entered the room in pairs and were asked to take place behind a table with a laptop. They were verbally asked to start the first game of the session. After this game a short evaluation took place in which participants were asked to reflect on the game and controls. After the evaluation the participants were asked to start the second game. They were not informed about the difference between the game conditions at forehand. After the second game the participants were informed that the game was made harder when GSR and heart rate increased. Before starting the third game condition, they were told that the game would become easier when the heart rate and GSR would increase. After the third game another evaluation took place in which the participants were asked to reflect on the changes between the game conditions, their game experience in the different conditions, on the controls, and what they thought about the mapping of the physiological data on the specific parameters.

RESULTS AND EVALUATION

First game session

In the first game session we observed that the users were in an uncomfortable position when they controlled their paddle. Players also indicated they would prefer a faster and better mapping of their physiological data on the game attributes. These results indicated that user experience could still be improved. Players were satisfied with the graphical user interface of the prototype.

During the first game session abnormal heart rate and GSR values were detected. Misplacement of the pulse sensor caused these abnormal values. The heart rate was dynamic and changed quickly. GSR values were not dynamic. In addition, GSR values were lower than expected due to a too low current caused by the use of a single breadboard powered by one Arduino. Since the data quality and execution of the prototype were not optimal in the first game session, the user performance of the first game session is not evaluated.

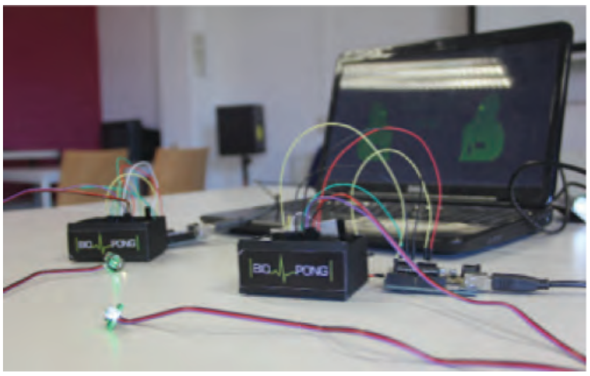

Figure 1. Controller used in the second game session.

Prototype improvements after game session one

To improve user experience, custom-built controllers were used instead of a keyboard. These controllers were connected to two different Arduinos with their own circuit board. The first prototype included only one Arduino and the additional Arduino in the second prototype increased the data transference speed from hardware to software. This led to a faster and better mapping of the physiological data on the game attributes. To improve the placement of the sensors an instruction screen was made to instruct the players how to place the sensors correctly.

Second game session

Figure 2 presents the quality of the data that was gathered from the sensors; the heart rate data, GSR data, and scores of two players in the second condition of the second game session.

During the second evaluation sessions players indicated that the instruction screen with sensor placement information was not clear. Other comments were made on how the controllers were working and looking. The potentiometer and button were too close to the other electronics and therefor not distinguishable. The test subjects indicated that they liked the games using physiological data (i.e., condition 2 and 3) more than the normal one. Participants were especially enthusiastic about the third game, in which the game became easier when the heart rate and GSR increased and harder when they decreased.

Scores of all players in the second game session are presented in Table 1. Analysis of variance was used to test whether mean player scores were significantly different across the three conditions. Player scores did not significantly differ across conditions. Thus, biofeedback did not improve player scores. Biofeedback also did not change a lot in ones performance against the other.

Prototype improvements after game session two

For the final prototype the instruction screen was redesigned. A case for the controller was designed with a 3D-printer. This case hides the electronics from the users and only shows the button and potentiometer. In addition, the button and potentiometer were increased in size and were given different colors to increase their visibility. The final prototype has not been tested yet.

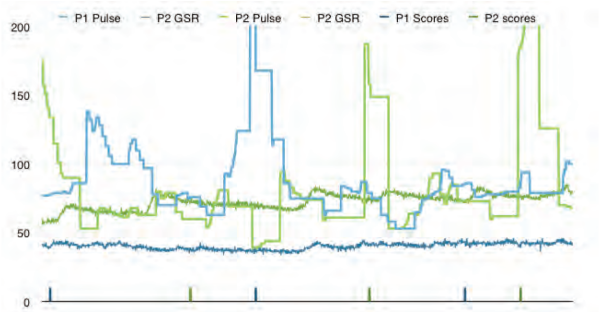

Figure 2. Heart rate, GSR, and scores of two players in the second condition of the second game session.

	Condition 1	Condition 2	Condition 3
Pair 1	7-7	11-6	4-8
Pair 2	4-7	6-11	7-10
Pair 3	2-2	3-3	3-1
Pair 4	5-6	10-7	8-4

Table 1. Player scores in the three different conditions of the second game session.

CONCLUSION AND DISCUSSION

The results of the present study suggest that physiological data can be used to improve user experience. User performance did not improve with biofeedback. In addition, since both player conditions were changed, gameplay modifications did not affect a lot in ones performance against the other. For future prototypes it may be interesting to test different mappings based on who leads in score. For example, the person who leads in score gets a harder gameplay and the other an easier.

The results indicated that the GSR values were rather stable. This means the use of GSR in a short and fast paced game like Pong is not suitable. The GSR may be a more suitable input for slower paced games.

The physiological state of the players was categorized in different ranges without taking into account base levels in heart rate and GSR of the players. Base levels differ between persons. This means that two persons that are in a similar emotional state have different physiological values. Thus, base levels should be taken into account to be able to say whether heart rate and GSR are elevated (indicating stress) or not. For example, a calibration can be made beforehand. Another possibility may be that the software learns the different emotional states of players the more they play.

Although the heart rate fluctuated a lot through game events, Pong may be not the best suitable game to evoke stress or other emotions due to the short duration of a game and its simplicity. Nevertheless, a game that lasts longer and in which the player experience more intense gameplay (e.g., for example first person shooters or sport games) may be more suitable for the use of physiological data to alter gameplay.

FUTURE RESEARCH

In this study two types of biometric data were mapped on two game parameters. Future studies can try to map physiological data to other attributes, for example the size of the ball and opacity of the paddles. In future research, differences between physiological states of the players can also be used to change their game attributes.

Although the sensors in the present study functioned well, higher performance sensors will improve the prototype with more accurate and stable readings. In addition, within the graphical user interface the difference between showing and not showing the physiological state of the user should be investigated.

The concept behind the prototype can easily be adapted to games with more competitive and complex gameplay where emotions rise higher and the game duration last longer. In addition, it can be tested on real life applications to improve users' ability to use interfaces in demanding situations. For example, by using the physiological state to alter the interface in order to compensate stressful situations, (e.g., applications for astronauts, soldiers, or surgeons).

ACKNOWLEDGMENTS

We thank Fons Verbeek and Jasper Scheffel for guiding and inspiring us during our project. We especially thank Rosanneke Emmen for her support. We also thank all the people who participated in our game sessions.

REFERENCES

1. Bersak, D., McDarby, G., Augenblick, N., McDarby, P., McDonnell, D., McDonald, B., & Karkun, R. (2001, September). Intelligent biofeedback using an immersive competitive environment. Paper at the Designing Ubiquitous Computing Games Workshop at UbiComp.
2. Dekker, A., & Champion, E. (2007). Please biofeed the zombies: enhancing the gameplay and display of a horror game using biofeedback. *Proc. of DiGRA*.
3. Gilleade, K., & Allanson, J. (2003). A toolkit for exploring affective interface adaptation in videogames. In *Proc. of HCI International 2003* (Vol. 2, pp. 370-374).
4. Gilleade, K., Dix, A., & Allanson, J. (2005). Affective videogames and modes of affective gaming: assist me, challenge me, emote me. *Proc. of DiGRA.*
5. Nenonen, V., Lindblad, A., Häkkinen, V., Laitinen, T., Jouhtio, M., & Hämäläinen, P. (2007, April). Using heart rate to control an interactive game. In *Proc. of the SIGCHI conference on Human factors in computing systems* (pp. 853-856). ACM.
6. Picard, R. W. (1999, August). Affective Computing for HCI. In *HCI (1)* (pp. 829-833).
7. Rouse, R. (2001). Games on the verge of a nervous breakdown: emotional content in computer games. *ACM SIGGRAPH Computer Graphics, 35*(1), 6-10.
8. Sykes, J., & Brown, S. (2003, April). Affective gaming: measuring emotion through the gamepad. In *CHI'03 extended abstracts on Human factors in computing systems* (pp. 732-733). ACM.

Digital Learning Environments and Collaborative Pedagogy: Media Culture 2020

Martyn Thayne
University of Lincoln
Brayford Pool
Lincoln
mthayne@lincoln.ac.uk

Graham Cooper
University of Lincoln
Brayford Pool
Lincoln
gcooper@lincoln.ac.uk

ABSTRACT

This paper presents an educational case study of *Media Culture 2020,* an EU Erasmus Intensive Programme (EU ERASMUS project number 2012-1-FI1-ERA10-09673) that utilised a range of social media platforms and interactive computer software to create open, virtual learning environments where students from different countries and fields could explore and learn together. The multi-disciplinary project featured five universities from across Europe – and was designed to develop new pedagogical frameworks that encourage collaborative approaches to teaching and learning. This paper will focus primarily on the implementation of a number of digital tools, in addition to highlighting the key educational aspects of the project.

Author Keywords

Convergence; Social Media; Collaboration; European Culture, Virtual Learning Environments

ACM Classification Keywords

H.5.3 - Computer-supported cooperative work; K.3.1 - Computer Uses in Education - Collaborative learning

INTRODUCTION

During the Spring and Autumn of 2013 five universities from across Europe took part in an innovative EU funded project, designed to explore how the integration of interactive digital technologies and social media platforms might foster new modes of collaborative teaching and learning. The principle objective of this project, entitled *Media Culture 2020* (‘*MC2020*’ hereafter), was to enable participants with a diverse range skills and cultural experiences to develop new working practices that respond to the convergence of digital media and art, as well as the internationalisation of media production and business.

The second main objective of the project was to break down classroom and campus walls by creating open, virtual learning environments where students from different countries and fields could explore and learn together. In short, *MC2020* was designed to interrogate the role of networked digital technologies in the development of pedagogy, demonstrating a number of ways in which Web 2.0 ‘architectures of participation’ [5] might be adopted by academics to encourage open and collaborative modes of practice. The project utilized a number of social media platforms (including Facebook, Twitter, Google+ Google Hangout, Google Docs and Blogger) to enhance the learning experiences of a diverse set of students from different cultural and international contexts.

The project was comprised of two two-week workshops, which both featured an additional 6 weeks of online activities, team meetings, interactive ‘webinars’ hosted by each partner university, as well as ongoing modes of social networking and collaborative practice. Added value was gained by these pre and post workshop activities through the implementation of ICT and social media services and tools. This involved the collaboration of students and staff members, as well as the involvement of other lecturers who could be part of the project virtually, without costs for travel and accommodation. The main activities during the pre-workshop phase were team building, project planning and researching online. The learning outcomes include skills in art and media production for 21st century platforms, market research, business planning, pitching, working in international, multidisciplinary teams and the application of social media services. The project outcomes included the production of a wiki - used for knowledge building - and the blog as a public-facing channel for exhibiting the new ideas and content created during the workshops. The blog and various social media platforms also enabled both staff and students to document the whole process.

METHODOLOGY

Media Culture 2020 was a multicultural project, featuring staff and students from five universities from across Europe: the University of Vic (Spain), Tampere University of Applied Sciences (Finland), Liepaja University (Latvia), the University of Lincoln (United Kingdom) and HKU Hilversum (Netherlands). The students involved were third and fourth year BA students of fine arts, interactive media, business, film and television, whilst the participating lecturers all came from different practice and theory backgrounds. The selection process took into account the level of English, the skills of students in collaborative work and in the use of new technologies. This EU Erasmus Intensive Programme was specifically designed to combine the diverse skills and cultural experiences of all involved to

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

develop new modes of collaborative pedagogy and digital scholarship.

The multidisciplinary approach utilised in this project is clear, with each partner university contributing different skills and knowledge to the project: Tampere offered expertise in interaction design and educational use of social media; HKU in applied narrative design and software & hardware development, University of Vic in entrepreneurship, business, audio visual and media production and blended learning, Liepaja University in combining virtual and physical worlds and immersive media, Lincoln University in games design, mobile phone gateway development, user experience design, convergent media practice, emergent media technologies and participative project development. The wide range of practice skills and research expertise made accessible through the international delegation of lecturers underpinned the whole experience, with students able to request feedback and advice on their project ideas, depending on their specific needs.

The two workshops, held in Tampere (featuring 10 lecturers and 40 students) and Liepaja (a further 49 students), were accompanied by a range of online pre and post workshop activities, with a series of seminars, group tasks and social networking extending and enhancing the teaching experience in a virtual learning environment. The focus of the workshops intended to interrogate the convergence of computer technology, media reception and art practice by exploring the potential of interactive media in the context of an increasingly multicultural European terrain. There were a number of seminars on the subject of interface design, ‘smart’ technologies, ‘open data’ and future developments in ICT, with groups having to explore these ideas, collaborating on mock-ups, workflow models, animations and concept designs. Whilst many of the students did not have much prior knowledge of these subjects, the multi-faceted learning environment encouraged more creative and open approaches to teaching and learning. Feedback and support was given by both staff and students, in addition to the dissemination of relevant research sources via social software. The implementation of the Facebook group page was useful in this respect, as it tended to result in less formal relationships between staff and students. As such, teaching and learning was encouraged outside the traditional parameters of the classroom.

In comparison to the courses that already existed in the partner institutions, *Media Culture 2020* utilised a more flexible and empowering educational framework for both students and lecturers. For the student, this approach generally led to an improvement of self-management; the implementation of collaborative work in a European environment; improvement of the quality of mentoring; and a diversification of activities and professional abilities. For the lecturer, *MC2020* represented an opportunity to partake in a new pedagogic relationship with students, which took the form of responsive, two-way dialogue; the implementation of a flexible monitoring and evaluation process where the two are blended together as one and the same; and through the diversification of tools for organizing activities related to content. Together, both staff and students could develop innovative practices related to digital collaborative work, engaging in diverse ICT and social media learning methods. Recent research [1];[4] indicates that when students work collaboratively in small groups they learn more and retain more, leading to a more satisfying and rewarding experience. Christopher McMorran [4] suggests that if used in an educational setting, collaborative technology can enhance active participation (through content creation), increase student engagement, and enrich the learning process. *MC2020* provides an exemplar of this model, utilising a range of collaborative technologies to produce a dynamic and democratic digital learning environment.

USE OF ONLINE TOOLS FOR COLLABORATIVE LEARNING

The central method for fostering collaborative practice was a mix of online learning and intensive workshop activities. The project presented an opportunity to develop a new kind of multicultural European mobility, using a range of ‘cloud-based’, social media tools to create joint virtual classrooms, labs and studios. *MC2020* utilised Google+ and associated applications (Google Docs, Google Drive and Google Hangout) as the core tools for the process. What makes Google Hangouts particular appropriate for this particular task was its ability to integrate Google Docs, screen-sharing and a streamlined ‘invitation to join’ process. Additionally, a ‘hangout’ can be saved to the YouTube platform for future referencing and even broadcast live. The use of these tools fostered a digital learning environment that extended the traditional boundaries of the classroom in time and space. Whilst these services are by no means unique comparison to other online tools [1];[6], we opted to use the Google software for a number of reasons. Most notably, we decided that the integration of a range of different technical features, coupled with the popularity of these services, made them an ideal choice to ensure that all participants had access to the same software. *MC2020* is not the first project to implement these technologies into teaching [5];[7], although the scale and multicultural dimension of this particular project does further highlight the value this approach.

When collaborative documents are prepared on GoogleDocs, there is only one version, which is always up to date and includes all corrections. This results in a more accurate and cohesive understanding of the project from all involved, since everybody has access to the same information. Whilst these features are perhaps common knowledge, they were particularly significant for this project in that the chosen service enabled ‘real time’ collaborative action between staff and students, regardless of geographic location. For *MC2020,* GoogleDocs was utilised due to its integration with a number of different software needs (word processing, spreadsheets and presentations) and with the ability to create these documents from scratch within the web browser. The associated ‘cloud’ storage service, GoogleDrive, allowed these documents to be shared instantaneously with students, whilst also facilitating a separate space for admin purposes. Throughout the project student groups each had their own folders for sharing work in progress, which the lecturers could also see and comment on if required. We even composed the initial proposal for *Media Culture 2020* using GoogleDocs, which was particularly beneficial in this instance since it enabled lecturers from each of the partner institutions to easily contribute to this document.

Alongside the Google services, *MC2020* utilised a number of social media platforms as a way of disseminating information and enhancing the relationships between students and staff from diverse cultural backgrounds. Whilst the volume of Google+ users has increased steadily, with 540 million active users as of October 2013 [3], it was evident that Facebook was a more widely recognised and utilised service amongst participants of this project. It was therefore decided that this more popular service be used as a social space, or virtual 'coffee room', functioning as an informal hub for sharing personal information and cultural exchange. It worked well as a platform for networking and interaction, and the groups and connections formed as part of this project still remain active to this day. It also had the added value of allowing staff and students to quickly share knowledge and pool relevant research sources, leading to further discussions and modes of learning in a more informal setting.

Due to the high profile of the project, funded by an external body, *MC2020* needed to present public facing content, disseminating proceedings and progress of the project. We also needed a service that would allow all lecturers to publish content. We opted to create a dedicated website/blog (mediaculture2020.blogspot.com), which once set up could be easily updated without the need for specialist computing knowledge. This site functioned as the curation of relevant information, presented in a formal capacity that distinguished it from the Facebook group. Although work was published to both, the context was informed by the delivery method and audience. One feature of the blog that was particularly useful was the inclusion of a Twitter plugin, which aggregated any information posted by staff and students to their own personal Twitter feeds via the hashtag '#MC2020'. Not only did this serve the purpose of publicising the project to a larger online audience, it also encouraged a more diverse documentation of the whole process, leading to a vast collection of associated tweets and images from the event. In the case of *MC2020,* then, microblogging services like Twitter and Facebook represented an excellent example of crowdsourcing, whilst simultaneously fostering more personal relationships between staff and students.

Pre-workshop Activities

The first phase of the pre-workshop activities took place in the six weeks (March 5 - April 12) leading up to the first workshop held in Tampere (April 15-28, 2013). During this period we split participants into five student teams comprised of members from each University. These teams worked online using the aforementioned online platforms to work collaboratively on three assignments. First, teams were asked to design a logo for *MC2020* and discuss issues of branding and visual style. The teams then had to choose two or three topics related to the project brief, researching these together and presenting a summary of their findings in Tampere. The final task set during this first pre-workshop phase was to make proposals to improve the draft programme of the actual workshop. These proposals were then voted upon, thus embodying the democratic approach we strived for throughout the project.

Pre-workshop activities for the second phase of *MC2020* took place during the first 4 weeks of October, which culminated in the second workshop in Liepaja (26 Oct - 8 Nov). The coordination team of the project had one preparatory meeting (via Google Hangout) in June, three in September and two in October. The pre-workshop activities included five online sessions, with lecturers from each of the partners' universities delivering an online seminar relevant to the project. Again, participating students were split into mixed-nation groups and worked on team projects during this pre-workshop stage. Using Google Hangout, Docs and other collaborative tools, the groups were asked to analyse one of the concepts developed during the Tampere workshop for further processing in Liepaja, with the results presented by the groups during the first working day in Latvia.

In theory, the idea of the pre-workshop phase of activities was a good one because it would engage students in both their local and international groups in order to build team bonds before the groups met 'in person' during the workshops. In the case of *MC2020*, this practice yielded some success. Whilst some activities struggled to attract engagement from all participants (especially those who were not overly confident in their ability to communicate in English), others exceeded expectation. This success suggests a model of remote, collaborative working could be pursued in other educational settings based on the approach taken.

CONCLUSIONS / RESULTS

In terms of collaborative working and the implementation of interactive, web-based technologies, *Media Culture 2020* was a success because the aforementioned barriers of remote working during the pre-workshop phase were overcome, whilst a more cohesive approach to sharing new knowledge was developed throughout. The various communication platforms utilised provided an appropriate toolset for documenting progress and experiences, in addition to facilitating a more open channel for the dissemination of information and feedback. Student evaluations of the project suggest this model of working provided more than just a set of tools to foster collaborative practice, it became a catalyst to change perceptions of trust and for enhancing bonds between staff and students.

For non-real-time collaborations the provisions of Google Drive and Docs worked as intended. It is evident these tools enabled all involved to have a consistent and seamless experience of contributing to tasks, fostering a strong culture of collaboration. However, as with the organisation of this project, activities that did require real-time collaboration proved difficult. This can somewhat be attributed to the fact that it is very difficult to assemble groups from multiple locations, time zones and schedules to be together and online at the same time. An example of real-time communication difficulties came in the form of shared lectures. Whilst the live streaming of these lectures was successful to some degree, the essential supporting visual content of some presentations, viewing them after compression and decompression for network transmission, failed to communicate some of the ideas presented for discussion. Despite these minor deficiencies in the approach we found that the project yielded a number of desirable outcomes:

- The open-ended brief and flexible teaching structure empowered students to define the working environment. The structure of the workshop themselves were open to negotiation, whilst students were encouraged to pool their collective research and practice skills.
- By leveraging the capabilities of Web 2.0 technologies this model of digital scholarship facilitated a more open, interactive and collaborative working environment for teaching and learning. Not only were students able to meet the formal assessed requirements of the project, they were also able to contribute to a wide range of intellectual discussions that were made accessible to all through the various software utilised. This can be seen as a more open process of learning since students were able to observe alternative ideas and work contributed by other participants, in addition to the collective feedback of staff.
- The technologies and virtual learning environments discussed above allowed for real-time collaboration whereby information and knowledge could be accessed and disseminated across a number of networked devices. This had particular value in the workshop phase by enabling students from different countries to work together.
- Peer-review and student driven feedback was given throughout the project. During the Liepaja workshop, final concepts were exhibited during iWeek, an international interactive arts event. A summary of feedback from this event was later published on the blog. Students were also rewarded with partial ECTS credits.
- Participants were engaged in active research activities throughout the conceptual development, presentation and delivery of projects. A collaborative approach to research was encouraged: we set up a 'library' of useful research sources, with contributions from both staff and students. This played out as a constantly evolving archive, connecting and pooling the research activities from both workshops.
- The established teacher/student divide was avoided wherever possible, with optional seminars, interactive workshops, student lead-presentations, group discussions and plenaries taking the place of the traditional, rigid lecture/seminar module structure.

The functionality of the Google software enabled students and staff from different cultural contexts to engage in a mode of collaborative learning that would have not been possible in the existing teaching infrastructure of each of the individual partner Universities. Whilst an ideal situation would be the development of a purely academic technological infrastructure that would permit much of the same modes of practice discussed in this paper, the use of Google software demonstrated a real benefit to this project. *MC2020*'s ambition to develop new collaborative approaches to pedagogy; to move from teacher-centred, standardised test oriented education to student centred, open learning focused education, was made possible my these chosen technical tools.

AKNOWLEDGMENTS

- *Media Culture 2020* - EU Erasmus Project Number: 2012-1-FI1-ERA10-09673
- Lead coordinators - Cai Melakosi, Tampere University of Applied Sciences
- MC2020 Blog - mediaculture2020.blogspot.com

REFERENCES

1. Berger, P. and Trexler, S. (2010), *Choosing Web 2.0 Tools for Learning and Teaching in a Digital World*, California: ABC-CLIO
2. Cunningham, W. (2002) '*What is a Wiki?*' http://www.wiki.org/wiki.cgi?WhatIsWiki (Accessed May 2013)
3. McGee, M. (2013), 'Google+ Hits 300 Million Active Monthly "In-Stream" Users, 540 Million Across Google', Marketing Land. http://marketingland.com/google-hits-300-million-active-monthly-in-stream-users-540-million-across-google-63354
4. McMorran, C (2013), Collaborative Learning Using Google Docs and Maps, *Technology in Pedagogy*, No15.
5. O'Reilly, T. (2005), "What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software," *O'Reilly Network* [online] @ http://oreilly.com/web2/archive/what-is-web-20.html
6. Solomon, G and Schrum L (2007), *Web 2.0: New Tools, New Schools*, Washington: ISTE.
7. Zink, A. (2012), "Google-Doc Surveys for Teaching Hispanic Culture" in Morgan, R. and Olivares (eds.) *Quick Hits for Teaching with Technology,* Indiana: Indiana University Press.

chi sparks 2014

The Kids' Knowledge Base: Connecting Junior Science to Society

Aldo de Moor
CommunitySense
Tilburg, the Netherlands
ademoor@communitysense.nl

ABSTRACT

Universities try to reinforce their connections with society in many different ways. Introducing children to science at an early age is an important part of this mission. The online "Kids' Knowledge Base" is a key instrument for presenting showcases of various scientific fields to primary school children, thereby aiming to pique their curiosity. We outline the architecture and development process of the Kids' Knowledge Base, and describe how it is increasingly being embedded in an ecosystem of online and physical tools, stakeholder networks, and activities. We show how it has been used since its launch in March 2013, and discuss how combining different modes of offline and online interaction helps to promote its overall usefulness and use. We discuss some applications and extensions of the current digital infrastructure and how these may help increase the quality and quantity of the online interactions with the knowledge base.

Author Keywords

Science; children; knowledge bases; e-learning; social media; socio-technical systems; social innovation

ACM Classification Keywords

H.5.3 [Information Interfaces and Presentation] Group and Organization Interfaces - Asynchronous Interaction

INTRODUCTION

Universities traditionally have been inward looking in their educational process, only teaching their mature students science. Increasingly, however, they aim to reach out to younger, pre-university populations. One reason for doing so is that being immersed at a young age helps the knowledge worker generation of the future to be more competitive [1]. In the Netherlands, there exists a national network of so-called "science hubs", each associated with one or more universities. The mission of these hubs is to find innovative ways to get primary school-aged children interested in science. The manifold activities of one of these hubs, the Wetenschapsknooppunt Brabant (Science Hub Brabant) include, for instance, "kids' lectures", in which professors at the university give special lectures about their topics of specialization to groups of children; "kids' science books", in which professional science journalists interview researchers about their projects, then retell their stories in language comprehensible to children (and many adults); and a "Junior Science Café", in which researchers discuss their work with children, jointly performing small experiments with them.

A major drawback of these physical activities is that they are expensive and hard to scale. Instead of just having local and print activities, a project was started to try and find a way to develop a digital resource expanding the palette of activities. This resource would need to be accessible everywhere by everybody at any time. It was dubbed the "Kids' Knowledge Base".

This short paper is not intended to provide a theoretical framework for human computer interaction in science communication, nor to describe a statistically sound experiment on how children interact with particular online media. Instead, our goal is to introduce a real-world educational case – the Kids' Knowledge Base - that is all about how such ICTs may come to have true societal impact. In our case, ICTs are a critical enabler, yet they do not mean much without equally developing a solid social context around these technologies. Ours is really a case of social innovation, where (socio-technical) innovations necessarily go through a spiraling process of scaling up from small-scale initial inspirations, ideas, and prototypes to large-scale systemic impacts via a series of intermediate stages [2]. Each stage is a mix of planned interventions and unpredictable events that offer opportunities for (and possibly threats to) further growth.

Storytelling is a powerful tool for communicating complex knowledge management and organizational change processes [3]. Although our case is still very much unfolding, we would already like to share our story of the interventions and events that happened so far and which helped to shape the Kids' Knowledge Base. Although anecdotal, we hope our story may inspire similar development projects elsewhere, as well as research into more systematic principles for developing such society-oriented, living educational knowledge bases.

KIDS' KNOWLEDGE BASE: THE DEVELOPMENT

The following objectives were adopted at the outset of the project. The Kids' Knowledge Base (KKB) should:

- Enthuse children for science.

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

- Develop scientifically valid content for curious children in primary school.
- Enliven content through activating conversational and work practices.
- Support and connect other science hub physical activities to the digital knowledge base.
- Make it a living knowledge base by developing a community of stakeholders around it, who continuously find new ways to use and feed the knowledge base and expand its network of applications and related activities.

Key stakeholders working directly with the KKB, besides the children, should be their teachers and parents (Fig.1). The teachers should be able to use the digital content of the KKB in class activities, while the parents should be able to process the same content with their children in one-on-one conversations at home.

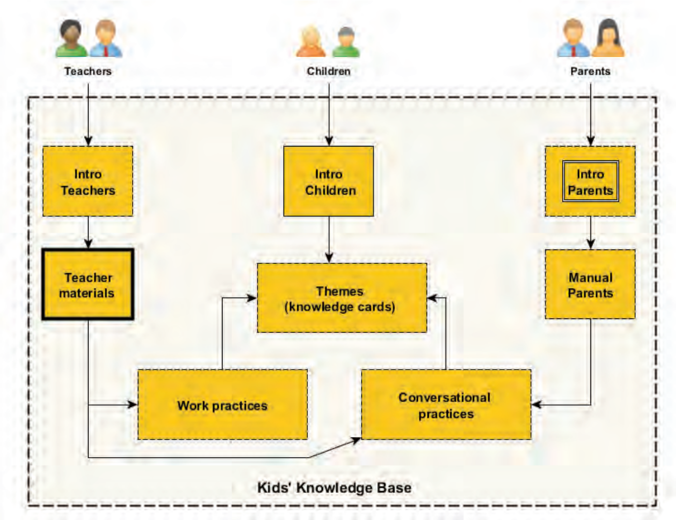

Figure 1. The Kids' Knowledge Base Architecture

The KKB contains a set of introductory knowledge modules ("kennismakingsmodules" in Dutch, which also has the meaning of "getting to know each other"), each introducing a different scientific field or area. Each module consists of a set of four or five themes, every theme in turn decomposed into a set of four or five "knowledge cards", explaining a particular topic belonging to that theme in a way comprehensible to primary school pupils. Each knowledge card has the following structure:

- Short introduction of the topic being presented
- Example
- Research related to the topic
- Research questions
- Advanced research (for very curious children)
- Related topics

Besides being organized by theme, knowledge cards also contain links to other cards, effectively creating a knowledge network. Of course, 20-25 cards cannot fully cover a scientific domain. In line with the project objectives, the purpose of a module is merely to pique a child's interest, then guide it to other (physical) science hub activities or more extensive digital materials, for example to be provided by libraries or publishers.

For parents, a simple manual suffices on how to have an informed discussion with their children using the digital knowledge cards. For teachers, a separate section was developed with additional open access learning materials, such as Powerpoint presentations that they can modify for their own classes, and tutorials explaining working and conversational practices to be used in promoting group discussions. At the teachers' request, this section was login-protected, to prevent children from downloading materials before class, so that the teacher can remain in control.

To produce a demonstrator of the KKB, a pilot was started to develop an introductory knowledge module for Philosophy. A teaching assistant (who was also a philosophy student) wrote the content for the cards, supervised by an established philosopher, so that the scientific value of the cards was guaranteed. Another teaching assistant then created the digital version of the cards, including providing them with a uniform layout.

TURNING A WEBLOG INTO A KID'S WINDOW ON SCIENCE

As the project team had only very limited development resources, it was decided to only use online platforms hosted in The Cloud. To develop the digital content, initially a Wikispaces wiki (http://wikispaces.com) was used. The advantage of this wiki was that it was easy to develop a content navigation structure, and jointly work on the rough content, including revision histories. However, for presentation and use purposes, this platform turned out to be less than satisfactory. A main drawback was the limited set of design options and templates, whereas a smooth presentation was key to enticing children to use it.

Development was then moved to a hosted WordPress site (Fig.2). Besides being an advanced content management system, WordPress sites are also standard-setting weblogs. Blogs are natural tools for promoting web learning [4]. A custom theme was modified, including a menu outlining the main themes of the knowledge module. On mouseover, the topic cards per theme become visible and clickable. The home page contains a photo carrousel with colorful pictures, each embedding an intriguing question that should get children interested. The initial color scheme was considered to be too boring by children participating in test panels. The colors were therefore made much brighter, with a purplish pink dominating – a color scheme loved by kids!

Figure 2. The Kids' Knowledge Base Home Page

Every knowledge card gets its own WordPress blog page. Each section of the page is preceded by an appealing icon, e.g. a colorful looking glass indicating the Research-section and a question mark indicating the Questions-section. Extensive use was made of YouTube videos, embedded in the blog page. A particular issue here was to look for fragments that are copyright-free, as several fragments initially selected turned out to be copyrighted Disney videos that were taken down by YouTube at some point.

The comments section that comes standard with each WordPress page was initially disabled. Although the team very much valued the feedback of children on the content, this would also mean actively monitoring and responding to those comments. As working with children is very sensitive, and no continuous capacity for moderating comments could be guaranteed, this option was switched off at first. However, a new project team member did not know about this policy and had activated the comments option. Since the feared and hard to manage spam-threads by kids "playing around" did not happen, we have left this option activated – for now.

Evaluation

Several versions of the site and learning materials were tested, in a classroom setting in a series of lessons, and in a lab test-session with four parent-child pairs. Each pair got their own PC. A 15 minute plenary introduction, was followed by a 30 minute free exploration of the KKB. The evaluation was concluded by a 15 minute plenary discussion. At those test sessions, each pair was monitored by an observer, logging their comments and responses as they explored the site. Some of these HCI observations, although anecdotal, are worth summarizing here, as they may inform similar projects:

Observations by children

"It's a nice way to learn a lot"

Many paragraphs contained too much text and too difficult words.

Some children skipped most of the text, others read everything. Some said that if they had been at home, they would have read everything.

The boys thought the original dominant pink color to be too "girlish", after which it was changed into a more purple hue.

The children all loved the YouTube videos.

Observations by parents

Texts should be short, fonts should be big.

Use more steps in navigating the scientific content.

Start with concrete content, only then introduce the more abstract concepts.

Use icons to indicate the (repeated) structure of the cards.

Should the parent guide the child (prepare beforehand) or follow the kid (as it follows its preferences)? Both routes should be supported.

Do away with distracting columns of other events, just focus on the content, otherwise you lose the child's attention.

An interesting (and in hindsight, predictable) clash between the classification-driven world of scientists and the "common-sense"-driven world of children was observed as none of the children in the test group would click the knowledge card which contained the word "teleology". When the title was replaced by "Does everything have a purpose?" it was much more palatable.

USING AND GROWING THE KKB

The KKB was launched in March 2013 (http://kinderkennisbank.nl), with a big launch event in the university auditorium. The event was discussed that same night at a provincial TV talkshow.

A cross-medial approach is used to promote the KKB, including a Facebook page and Twitter account. Increasingly, the other activities of the Science Hub, such as the Junior Science Café, start pointing to it and using it, for example, in preparation for an event.

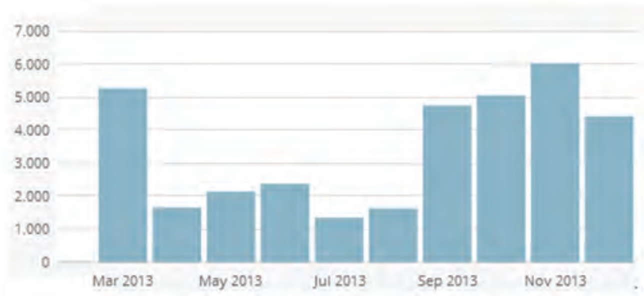

Figure 3. The Kids' Knowledge Base Page Views in 2013

Page views are surprisingly high and continue to grow, even though there is still only knowledge module (Philosophy) and there are no major PR efforts to promote the KKB. As of the writing of this article (March 10, 2014), there have been 54,272 page views, the highest ever being 1,271 page views, on average 217 per day. Unique visitors are also high, for example, November 2013, the peak month that year, counted 6,018 views and 2,728 unique visitors.

When looking at the monthly statistics, there is a significant increase of views in September 2013. This may be explained by a very interesting associated event held then, the "B@ttleweters Kids' Knowledge Battle". B@ttleweters (freely translated as "B@ttle Know-It-Alls"), was successfully launched during the European Social Innovation Week in Tilburg in September 2013. The format was developed by the provincial library innovation organization Cubiss, in association with Wetenschapsknooppunt Brabant. The knowledge battle combines the KKB with public library materials, services, and locations. Groups of children from participating schools prepare for the battle by working with the KKB in the classroom. In this way, they are taught both hands-on introductory science and media literacy skills. During the live battle, they go online to look up answers to questions derived by librarians from the KKB, while using a public library "media bar" to get access to networked devices. They also debate a proposition in front of the audience. The children are wildy enthusiastic, as this evidence shows: http://battleweters.nl/

B@tlleweters is currently being developed into a full-fledged kids' knowledge battle format, for which interest has been shown by libraries and science hubs all over the country. This is a good example of scaling up the partner network around the KKB, towards reaching social innovation levels of "systemic impact". By embedding quite basic digital resources in a well thought-through context of physical activities, the quantity and quality of the interactions – and their impact on promoting science to kids - can thus be amplified significantly.

New knowledge modules are currently being developed, such as around the European Values Study being coordinated by Tilburg University (which has been monitoring the evolution of social norms and values in many European countries over the past decades) and one on emotions. For the latter, content from an existing "children's science book" will be repurposed by the (science journalist) author, and relevant fragments turned into paragraphs for knowledge cards. To make the experience more interactive, we are currently exploring the possibilities of an established interactive quiz platform (http://www.proprofs.com/). Key requirements are that children can take the quizzes anonymously and that their metadata do not contain personal information.

CONCLUSION

As Jackson states, universities are above all social institutions that are "an essential part of the fabric of a vigorous and dynamic civil society, both contributing to the wider life of that society and at the same time open to the impulses and energies flowing from that wider life [6, p.105]". Universities reaching out to children and getting them interested in science at a young age is an important role to play for such a social institution in the emerging Knowledge Society. We presented how the online Kids' Knowledge Base is becoming a relevant instrument for science education. We do so by developing online introductory scientific content, activating this content through applications like B@ttleweters, and developing a strong partner network of stakeholders.

Although the digital tools used are essential to enable the many (potential) interactions, the tools are only a small piece of the puzzle. A key question is how to embed these tools in a complex, evolving socio-technical context driving the use and evolution of the platform. In this way, the Kids' Knowledge Base is becoming a truly living socio-technical system, impacting society in increasingly powerful and surprising ways.

ACKNOWLEDGMENTS

The author would like to thank all members of the project team and the Kinderkennisbank Partner Network, in particular Cubiss, for their many contributions.

REFERENCES

1. National Research Council. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Washington, DC: The National Academies Press, 2012
2. Murray, R., Caulier-Grice, J., & Mulgan, G. (2010). *The Open Book of Social Innovation*. London: NESTA & Young Foundation.
3. Denning, S. (2001). *The Springboard: How Storytelling Ignites Action in Knowledge-era Organizations*. London: Routledge.
4. Efimova, L., & Fiedler, S. (2004). Learning Webs: Learning in Weblog Networks. In *Proc. of the IADIS International Conference on Web Based Communities* 2004, Lisbon, Portugal (pp. 490–494). IADIS Press.
5. de Moor, A. and van den Assem, R. (2013). Public Libraries as Social Innovation Catalysts. In *Proc. of the* 10th Prato CIRN Conference "Nexus, Confluence, and Difference: Community Archives meets Community Informatics", Prato, Italy, Oct 28-30 2013.
6. Jackson, R. (1999). The Universities, Government and Society. In D. Smith & A. K. Langslow (Eds.), The Idea of a University (pp. 91–105). London: Jessica Kingsley Publishers.

GETTING IT DONE

‘Include’, a Toolbox of User Research for Inclusive Design

Hanna Zoon
Eindhoven Technical University
Department of Industrial Design
P.O. Box 513, 5600 MB Eindhoven,
The Netherlands
info@hannazoon.com

Anita Cremers
TNO
P.O. Box 23, 3769 ZG Soesterberg,
The Netherlands
a.cremers@tno.nl

Berry Eggen
Eindhoven Technical University
Department of Industrial Design
P.O. Box 513, 5600 MB Eindhoven,
The Netherlands
j.h.eggen@tue.nl

ABSTRACT

In order to empower more people to become more self-reliant in society, interactive products and services should better match the skills and values of diverse user groups. In inclusive design, relevant end-user groups are involved early on and throughout the design and development process, leading to a better user experience. However, for IT businesses not operating in the academic domain, getting access to appropriate user research methods is difficult. This paper describes the design and prototype development of the Include Toolbox, in close cooperation with practitioners of small to medium sized enterprises (SMEs) in IT. It consists of an interactive app paired with a book. The app helps to find suitable research methods for diverse user groups such as older people, people with low literacy, and children. The book offers background information on the advantages of inclusive design, information on different user groups, and best practices shared by other companies.

Author Keywords

Inclusive design; user research methods; low literacy; elderly; children; SMEs; IT; societal participation

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION

The world changes at an amazing pace, and technological innovations happen faster and faster [1]. Behaviour, needs, wants and likes of people also change, and many user groups are less homogenous than before; fewer and fewer people can be characterised by a traditional stereotype [32]. Older people: a 70-year old internet pioneer taking programming classes, and a 90-year old going shopping by bicycle. Disabled people: a 16 year old teen who battles BMX riders in the skatepark and calls himself an extreme sportsman, but also happens to be in a wheelchair [35]. Businesses cannot rely on existing knowledge alone; they need increasingly up-to-date insights into the perception and behaviour of their diverse customers, and need inspiration on how to shape the future of their enterprise.

As public services and products become more interactive and are increasingly offered solely online, it is important to not exclude anyone from using these services. In many countries, there is already legislation to ensure access to essential online services and products for everyone [34], for example ensuring a blind person can buy a train ticket, or a low literate person can apply for benefits online. But also for less essential interactive services and products, it can be an advantage to design them in an inclusive way, taking into account the diversity of the user group.

Inclusive design can be defined as the design of mainstream products or services that are accessible to, and usable by as many people as reasonably possible [20]. In inclusive design, relevant end-user groups are involved early on [30] and throughout the design and development process [3], leading to more varied inspiration and more user-friendly solutions. Making a product more inclusive, usually means better usability for everyone [7]. It is not only a way to solve problems, but also a strategy to identify problems to solve. With the growing importance of brand image and sustainable business practice, inclusive design can be a selling point and may broaden the potential customer base.

It can seem easier and cheaper to self-reference than to involve users. But to cater an application to a 28-year-old male programmer, is actually designing for a tiny minority. It can be very costly to launch an application and only then discover it does not suit the people who were intended to use it. From our previous research, it can be concluded that awareness of inclusive design is still lacking. Unfamiliarity with its advantages and methods seems to be a barrier to practicing it.

User involvement does not need to be difficult, time-consuming or expensive. For a business seeking inspiration from customers, or wanting to get a feel of the perception of their product by users, smaller informal sessions suit an iterative design process much better and are more practical. Sometimes, even one participant can be enough to get inspiration for new business ideas [27], and a session with five participants can uncover most of the general attitudes about a subject [6]. User research methods are generally very robust and even shorter, simplified execution will give usable results.

It can be challenging to find a user involvement method suitable for the goal of a project, time frame, resources and the type of target users. Academic as well as popular literature describes hundreds of research, ethnography and usability methods [Sanders, E.B.N., personal communica-

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

tion], [4]. Especially for businesses not operating in the academic domain, getting access to and finding the right research methods can be a challenging task, that is often not even started. As a result, users are often not included in the development process.

There are many websites and toolkits available that offer techniques to help getting new inspiration by thinking differently, and tools for visual design. There are only a few that offer user involvement methods, or are designed for use by practitioners in the information technology domain (IT). In this paper, we will describe the design and prototype development of a toolbox for inclusive design and end-user involvement, during which we worked closely together with practitioners of small to medium sized enterprises (SMEs) in the field of IT, that had little or no prior experience with user research.

The Include Toolbox prototype [26] that resulted from this cooperation consists of an interactive app paired with a book. The Include App (Figure 1) helps to find suitable research methods for diverse user groups such as older people, people with low literacy, and children. The Include Book offers background information on the advantages of Inclusive Design, information on different user groups, and describes best practices shared by other companies.

The Include Toolbox can help bridge the knowledge gap between academia and practice, and make Inclusive Design possible for more businesses. Ultimately, the toolbox could lead to interactive products and services that match the skills and values of diverse target groups better, in order to empower these people to become more self-reliant in society.

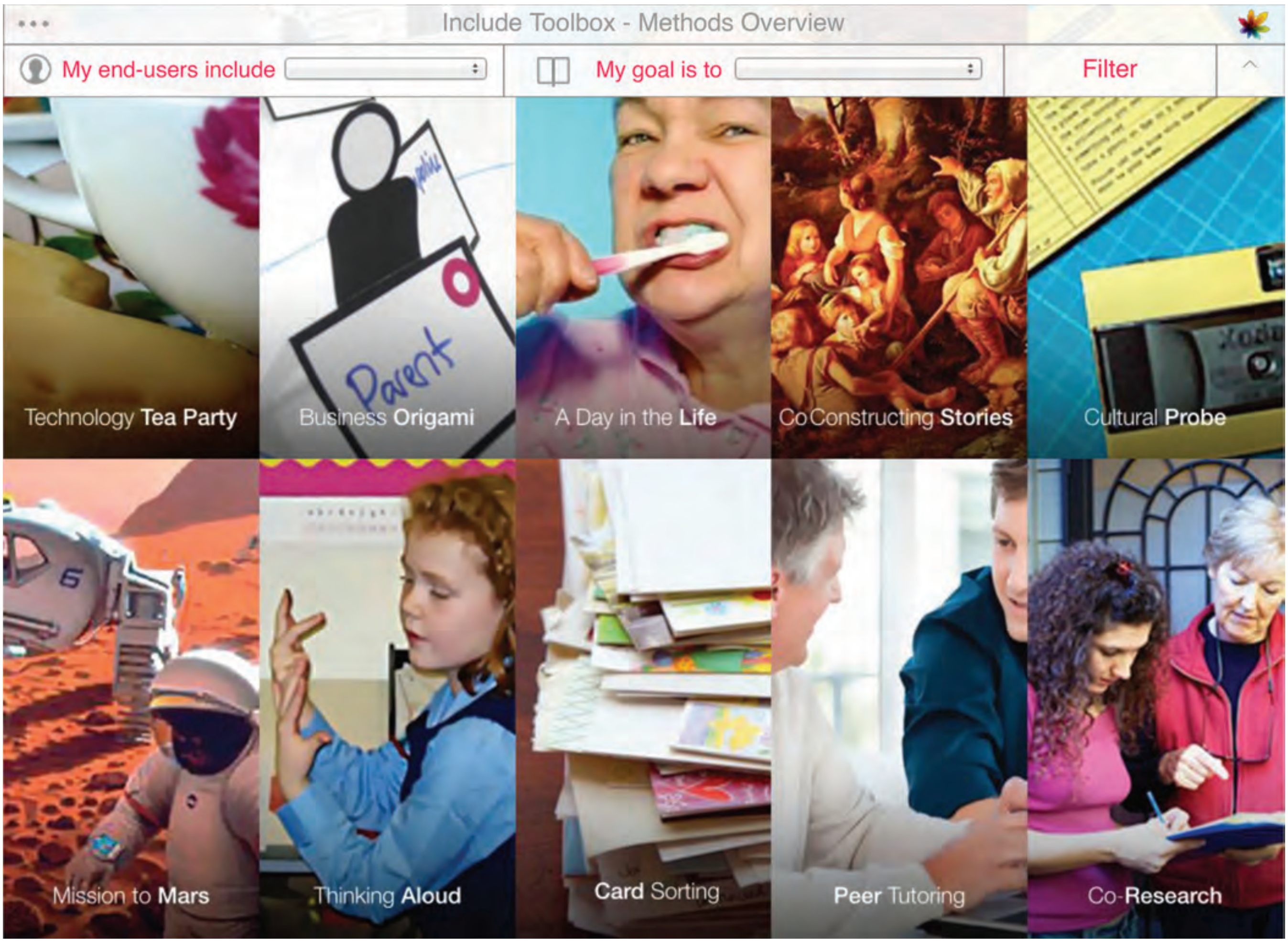

Figure 1. Home page of the Include Toolbox App (below)

RELATED WORK

Amongst the many toolkits available for designers, most offer inspiration techniques and tools for visual or product design. There are only a few that offer user involvement methods, of which a selection is described here.

IDEO cards [17]

The classic IDEO method cards are widely used and the methods on them became accepted practice in a lot of businesses. The deck of 51 sturdy cards, developed by the IDEO design firm, describes mostly inspirational methods and also some user involvement methods, and are specifically geared towards product designers. Already over 10 years old, it is one of the most well-known collections of methods and is often used as basis for other card sets and toolboxes, such as HCD Connect and Designing with People (see below). On the cards is a colour-coded system that helps designers to pick a suitable method. Of the toolkits described here, it is the only one that is not free.

HCD Connect - Bill & Melinda Gates Foundation / IDEO [14]

The Human-Centered Design Toolkit was designed specifically for people, nonprofits, and social enterprises that work with low-income communities throughout the world. The HCD Toolkit walks users through a human-centred design process and offers design- as well as user involvement methods, bases on the IDEO cards. It is divided in three parts: Hear, Create, Deliver; with the appropriate methods listed under pictograms.

Designing with People - Helen Hamlyn Centre for Design, Royal College of Art [11]

Designing with people is a website with very complete information on research methods, design methods, user groups background and stories, personas, and tips on doing user research. It is set up in a clear way, so that it is easy to decide which part of the information to view. The research methods that are available in this toolbox, are based on the IDEO cards and described in a compact way but further references are also given. The Designing with People website is one of the outputs from i~design 3, the final phase of a collaborative research programme on inclusive design funded by the Engineering and Physical Sciences Research Council. It has been designed to work in conjunction with the Inclusive Design Toolkit developed by project partner, the Engineering Design Centre at Cambridge University.

Inclusive Design Toolkit - Engineering Design Centre, Cambridge University [15]

The Inclusive Design Toolkit, next to explaining inclusive design and offering business rationale and design patterns, offers physical tools to assist inclusive design, including a Vision and Hearing Impairment Simulator, Cambridge Simulation Gloves and Glasses, and an Exclusion calculator. The design patterns mention user involvement methods but for descriptions refers to the closely linked Designing with People website. The Engineering Design Centre also offers courses on inclusive design.

UCD Toolbox - Tristan Weevers UCD / Delft University of Technology [12]

The UCD Toolbox is a very complete resource of methods for user centred design. It contains an overview of 35 design methods, which can be filtered by criteria: type of product, design goal, resources, participants and method characteristics. Also, a pre-selection of methods can be made for various target groups: elderly, children, physically challenged, visual/hearing impaired or cognitively challenged. However, no background information on specific target groups is offered and it does not become clear why the methods are suitable for the target groups. Method descriptions contain a lot of information and it is not always clear how much time, skill and effort a method will take. Because of the extensiveness of this toolbox, it can be hard for IT practitioners to find the specific information they need.

UX Toolbox: Better Web for Citizens - British Columbia Government [17]

A resource for government interaction and web designers, the UX Toolbox is a complete manual containing information on user experience, design research, web strategy, information architecture, content design, and web standards. The design research section describes research methods in detail, some of which involve end-users. It includes, research plans, reporting and managing tips. The methods are sorted into categories based on the type of research method, of which the terminology can be hard to understand for non-experienced researchers. The UX Toolbox also has a team that practitioners can contact for information and help.

55plusToolbox – Saxion [13]

The 55plusToolbox (in Dutch) focuses on topics that change the innovation process as a result of choosing the user group of people over the age of 55. Target users of the toolbox are entrepreneurs, focusing on both product development and marketing and sales. It contains information on the user group as well as case studies. Suitable tools for the particular phase and user group are suggested and illustrated in factsheets containing step by step guidance, visualisations, relevant links and references.

Universal Methods of Design - Bruce Hanington, Bella Martin [10]

The fact that Universal Methods of Design is a physical book, allows for another kind of interaction. When flipping through the pages, one gets a clear overview but not too much information at once, even though the book describes a total of 100 methods. The book contains design-, research- and user involvement methods, ordered alphabetically and coded for each design phase. The target users of the book are designers, but it is written in a very accessible way. There is a short description given for each method, completed with references and examples.

All toolkits and related work mentioned above are in theory also usable for IT practitioners. However, because of factors such as terminology and emphasis on creative techniques, some are clearly meant to be used by designers, not programmers. The British Columbia UX Toolbox is made for the IT-practitioner target group, but requires previous knowledge of and experience with user research. In most existing toolboxes, it can be difficult to find a suitable method, sometimes because there are so many, sometimes because it is difficult to get an overview [20]. Often exact descriptions of a method are not given, sometimes there are references to this information. But even then it is often still hard to determine the level of difficulty of the method, and the amount of time it will take. Few toolboxes offer sufficient background information on the needs and abilities of different target groups. Many of the resources mentioned above, cover some aspect of what we think an inclusive design toolbox needs to offer, but there is not yet a resource available that caters specifically to the needs and wishes of practitioners of SMEs in IT.

REQUIREMENTS TOOLBOX

In earlier research [3], a workshop setting was used to gather preliminary requirements for an inclusive design toolbox, which formed the basis for this project.

To further define and refine the requirements, we worked together with a total of 11 practitioners from seven SMEs in IT. The project was set up in an Agile way [18] with many short iterations. In the earlier stages of the project, focused interviews were conducted about existing business processes and views on inclusive design. In later stages, two practitioners worked with the toolbox in their projects [19] (Figure 2), providing feedback on the practical applicability, while others provided feedback on design and business aspects. Throughout the project, new companies were introduced to get fresh insights from people who had never seen the toolbox before. Five companies were smaller (5-20 employees) privately owned businesses, one practitioner was a PhD. student in the field of computer science and one company was medium-sized (~200 employees), all based in the Netherlands. The business activities included designing, developing and building websites, intranets, e-learning applications and apps, and consultancy.

In the beginning of the project, a very persistent attitude we saw in the companies was that user research would be difficult, expensive and not useful. Statements such as ‘users do not know what they want’ and ‘they never say anything useful’ were heard. With further questioning, it became clear that the people we interviewed were more frustrated with the lack of success, than convinced of the uselessness of user research. For example, they hated it when their app for helping elderly using public transit came on the market but was rejected by the target group, who claimed they did not need help. The company was now wondering if working with older people in a more structured way from the beginning, could have prevented the need for an expensive redesign. Another business owner saw the business opportunities of making a clients website more inclusive for elderly, attracting more customers that way.

Figure 2. Practitioner performing user research using the Business Origami method from the Include Toolbox

By working closely together with practitioners and owners from these and other companies, and using the knowledge gathered in our earlier research, requirements were refined and tested through prototyping throughout the project. This process allowed for a thorough understanding of the needs and business practice of our target user group, SMEs in IT. It also made it possible to adapt and change the design and setup of the toolbox continuously, to suit the requirements more optimally.

The requirements we found for an Inclusive Design Toolbox for SMEs in IT:

1. Easy to find the right method

Especially because it is so difficult for most people, and therefore also end-users, to verbalise needs and imagine products or services that do not yet exist, it helps to employ a structured process for this task [25]. For academics, searching a scientific literature library is easy. Finding a specific, suitable method without knowing its name is already more difficult. For non-academics there are books and websites on user research, but because this is such a big field, the amount of information is typically very large, and the information is often geared towards experienced researchers, not IT practitioners. The practitioners from our interviews were often not aware of the existence of scientific libraries or that there are many different user research methods.

2. Decision aid versus freedom of choice

When looking for suitable user research methods, practitioners wanted the autonomy to explore the methods themselves first, to see what there was on offer. Only after that, they would request toolbox recommendations. It was mentioned many times that they would like to keep their options open and choose themselves, and have the freedom to deviate from recommended methods.

3. Methods: fast, robust and cheap

For a scientific research project, it can make perfect sense to do a three-week long experiment with 120 participants. But in the IT world, with Agile development cycles, the time or budgets for this are often simply not there. And as Nielsen [23] and Dix [6] argued, with about 3 to 5 people, you will see most of the common behaviours from an entire group, and each extra participant gives only limited extra insights. Limiting the amount of participants limits the expense of time and money, making user research more likely to take place.

The methods should be robust so that even inexperienced researchers making mistakes, can get useful information from them. Therefore methods should be less dependent on facilitator skills, and require only limited previous knowledge of the target groups. Since sometimes the people in IT businesses who want to do user research, have to convince others, or need some convincing themselves, the methods that are offered must be well-known or validated methods, to inspire credibility.

4. Background information

Once the practitioners became more convinced of the importance of taking the needs of end-users into account in the development process, they often recommended that the toolbox would have ready-made personas and background information on user groups.

5. Examples (best practices)

Another requirement that came out of the interviews, was to supply examples or best practices, so that the practitioners could see how other businesses had approached user research. This way, it would also be possible to give examples of the type of knowledge one could expect to gain from user research, without presenting it as if this particular knowledge is generally applicable.

6. Business processes

Amongst the companies we worked with, there was a range of business processes practiced. Practitioners who were used to Agile or Lean [28] ways of working, were often looking for ways to verify early concepts with users within a short timeframe, because these methodologies require new versions or changes to a product to be validated of before the next iteration can begin. In more traditional linear (waterfall) processes, there appeared to be a bigger concern with the cost and validity of user research in general. If a project went over budget, evaluations were one of the first things to be skipped to save costs. User research was often only done when they needed specific proof to back up decisions. But even here, there was a general acceptance of the idea that it is cheaper to identify issues sooner rather than later.

7. Brand image / Commercial value

Next to the user research activities, it seemed important to companies to be able to show these activities off, and be able to use inclusive design as a sales tool. While an app or online toolbox would be practical in use, a physical component that can be shown in the office or given to clients is also important.

Being able to talk about inclusive design as a complete strategy, rather than separate user research methods, can show a brands involvement in sustainable business practices. It is important that the toolbox allows IT practitioners to learn about inclusive design as a holistic concept, to be able to sell it convincingly.

8. Terminology and tone

Some of the practitioners we worked with, joked about the words ‘design’ and ‘research’ as concepts they were not interested in. They saw themselves first as developers, as clients often provided the visual design and not much research was done. To avoid estranging practitioners with design jargon or academic terms, it is important to use more generic language or use terms from the IT world. On the other hand, the people that we worked with were highly educated entrepreneurs and independent thinkers, who did not appreciate to be talked down to.

DESCRIBING THE INCLUDE TOOLBOX

For a small to medium-sized company that works in the field of IT, ‘Include’ is a toolbox of user research methods providing an easy and efficient starting point for inclusive design. The toolbox is an interactive application accompanied by a hardcover book. The interactive part gives an overview of easy user research methods, complete with descriptions and workbooks, and helps filter them according to the company’s needs and their user groups. The book serves as a visual reminder, a reference for background information and a way to show off the use of inclusive design practice in the office.

The toolbox is currently available as a free application in Beta, that has the purpose of creating awareness about inclusive design in the IT sector. For future versions of the toolbox, different business models could be considered. The main functionalities and design elements of the toolbox are described below and structured according to the requirements listed above.

1. Directly from the home screen of the app, it is easy to find the right method

Incorporated in the interactive toolbox are currently 10 user research methods. This number of methods was arrived at by weighing different factors: there must be enough methods to suit four development phases and three specific target groups, but not so many that it would become overwhelming, especially for first-time users. Some methods are suitable for more than one phase or user group. The methods are all existing, validated methods, with references. The methods can be browsed from the main screen of the app, where after clicking a short description is offered next to a recognisable picture for each method (Figure 3). There is a filter bar, that can be hidden when not in use, that contains two drop-down menus: “My end-users include: [Older People, Children, People with Low Literacy]” and “My goal is to: [Get Inspiration, Evaluate an Idea, Review a Scenario, Test a Prototype]”. Using the filter creates a recommendation of three selected methods, that can be browsed through before making a choice.

Figure 3. Example of a method description card

2. The toolbox can act as a decision aid but also allows for making alternative choices

Offered are two ways of looking for a method (browsing and filtering). This has the benefit of being able to give people who know what they want direct access to the methods, while users looking for guidance can get it while keeping the option to choose open.

3. The methods are fast, robust and cheap

From the requirements description, we distilled the following criteria for user involvement methods that could be useful to SMEs in IT:

- it must be possible to do the preparation, session and analysis in a total of 8 hours, one half day at the beginning of the week and another at the end;
- no special skills or prior experience are necessary;
- it must be a well-known or validated, existing method;
- it looks like a fun activity for both the practitioner and the participants;
- all the necessary materials can be provided in a PDF workbook in the form of tips or worksheets;
- get useful results with 3 to 5 participants.

From a literature study, the following methods were selected based on the aforementioned criteria:

A day in the life [9]	Cultural Probe [8]
Business Origami [10]	Mission from Mars [5]
Card Sorting [31]	Peer Tutoring [21]
Co-Constructing Stories [25]	Co-research [33]
Technology Tea Party [2]	Think Aloud [22]

The methods are described in the following way: first a short introduction, on the basis of which the method is chosen. Then there are seven cards used within the app:

1. overview of the method;
2. description of the preparation, step-by-step, but short (Figure 4);
3. guide for the session itself, idem;
4. tips for a quick and effective analysis, idem;
5. filling in the name of the project and a short description;
6. choosing how to use the PDF workbook: print or digital, and after the research, archiving the work;
7. rating the method on suitability for the user group and ease of use of the method itself, providing tips and comments, choosing to share or not.

4. and 5. Background information and best practices are available in-app and in a hardcover book

User group characteristics change, and due to market fragmentation, user groups are not as homogenous as before [32]. Providing pre-made personas (in contrast to carefully constructed personas based on relevant user research data) would mean to overly simplify and stereotype the people of a user group [24], ignoring the needs and wants of sub-groups or failing to spot underlying patterns in a user group specific to a certain project. Therefore, even though practitioners asked for ready-made personas, we decided against this.

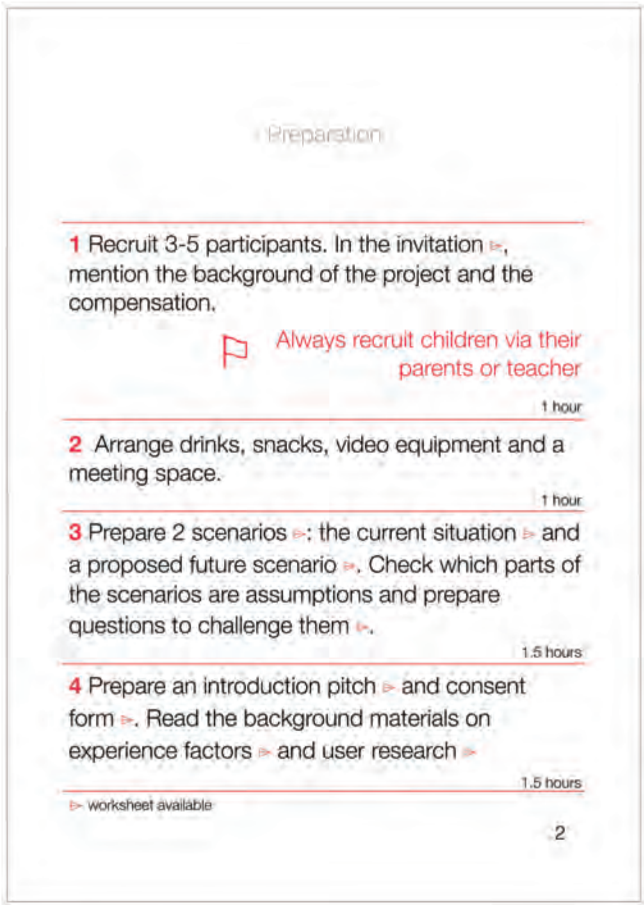

Figure 4. Example of a method workbook card

The toolbox does provide general background information about the larger target groups, to get a first idea of the characteristics of the group, and prepare the user involvement sessions more efficiently. In the app, a menu can be shown for access to more in-depth background information, such as about inclusive design, the methods, user group information and best practices.

From the method descriptions, there are also links to this information. Because it is unlikely that users are going to read many pages from within an app, the complete information is also available in the form of a hardcover and e-book. The contents of the book and the app overlap, to accommodate both easy reading and flipping through paper pages, as well as the ease-of-use, portability and the help of links within in the app.

6. Implementation in iterative business processes

Because of the limited preparation and analysis time needed for the methods covered by the toolbox, they are suitable to be implemented in business processes with short iterations such as Agile or Lean. User centred design based on design cycles is also a good fit. For other project structures, early user research can still be beneficial to starting the project on a relevant course, and also wherever there is a decision moment. In some cases, using the toolbox may even help practitioners to see their development process clearer, because the preparations for user research often include describing and focusing on the end-goals of the project.

7. Brand image / Commercial value

Research results can be used as proof to clients, validating development decisions. Being able to show that their company is practicing inclusive design was also perceived to be a selling point towards new clients. The book accompanying the interactive app allows for that, and is also a visual reminder in the office to think about special user groups more often.

Another aspect is that being able to show customers what user research is, can convince them a project needs it. It also makes it easier to talk about inclusive design, if clients can learn about it from a third party. More efficient development processes combined with a better brand image, can increase the overall competitiveness of a business.

8. Terminology and tone

The descriptions of the methods are void of design jargon and academic terminology as much as possible. The background information in both the book and the app, is written from a perspective of an SME entrepreneur, or IT practitioner, linking back to what they can get out of it in their work, or what it will do for their businesses. Furthermore, the steps of each method are balanced so that they give just the necessary amount of information without overwhelming or condescending.

EVALUATION OF TOOLBOX

In addition to working with different SMEs in IT throughout the project, the final prototype of the interactive toolbox app [26] was evaluated qualitatively on visual and interaction design and user perception with five participants, in three sessions. The participants of these sessions were practitioners in IT, with limited previous experience with user research. The evaluation method used was based on the Co-Constructing Stories [25] interview technique.

All participants said they thought they could easily execute the methods offered by the Include Toolbox, and appreciated the detailed workbooks. In general they were surprised each method could be executed in 8 hours and realising this was often a pivot point in the interview; from talking about difficulty and hurdles to talking about the benefits of involving users. They agreed that the toolbox language should be available in Dutch for ease of use with end-users from the Netherlands, especially for the target groups older people, people with low literacy, and children, who may not speak English. Another overall feedback was that the participants could appreciate the visual language of the toolbox, which was seen as colourful and inviting, and the information architecture, which was perceived as easy to use.

Participants commented on how they would use the toolbox app, by saying they would browse the overview page for a while first, before choosing a method to work with. Methods with ‚research' in the name came across as boring and formal, and some inspirational methods were seen as too creative. Getting a recommendation of three selected methods when using the filters was a good number for participants. In practice, participants would like to spend a minimum amount of time in the app, just quickly choose a method, print everything out and get to work.

Some participants would have liked to see more best practices, some said they would need much more different user groups and more specific user group information because for each project the target user is different. At the time this evaluation took place, the content of the toolbox prototype was not yet complete, and some participants were eager to get the remaining method descriptions, perhaps showing their motivation to get started with the methods.

Participants also noted that it seemed to them that answering the questions in the filter bar of the app, would already bring some awareness to practitioners about inclusive design. Using the toolbox to educate their clients about user research, could give the clients more confidence to discuss the details of user research, making them more satisfied with the company in the long run. There were some concerns if the option to share the research result as a best practice, would safeguard the anonymity of the end-users enough.

The hardcover book that accompanies the Include App (Figure 5), was not yet finished by the time the final evaluations took place. During the sessions, an interesting contradiction arose when discussing background information about topics such as inclusive design, user involvement methods, user group descriptions and best practices. For many practitioners, it was important to know it was offered and how they could reach it easily. However, when it was visually represented in the toolbox, links to the information were appreciated, but not opened. In an informal evaluation in an exhibition setting that did show the book, people were mainly attracted to the big size, and colourful layout.

Figure 5. The Include Book

CONCLUSIONS & DISCUSSION

Even though there were some criticisms from the participants in the final evaluations, the overall feedback was positive, with participants envisioning their use of the Include Toolbox and expressing motivation to use it in their work. The toolbox app could perhaps reach more practitioners if it allows for more company profiles, next to the currently very specific: practitioners from SMEs in IT with little or no experience in user research.

Some participants would have liked to see more best practices in the toolbox, and this is indeed planned by allowing users of the toolbox to share their own experiences. This way, over time the toolbox content becomes an increasingly complete reference guide. Further progress towards a more complete and continuously updated toolbox could be made through making it possible for practitioners to add or edit information, for example add new methods, update target group information, provide tips and rate methods.

Other practitioners wanted more specific user group information, but by their own admission the specifics are unique to each project. It is our intention to add support for more end-user groups such as teenagers, non-native people, or people with physical or mental disabilities. In our opinion, next to using the general information in the user group descriptions, it would also be necessary to work closely with users in each project, thereby making sure of having the most specific, unique and up-to-date insights as possible.

In our evaluations where practitioners used a prototype of the toolbox in real projects, we and the practitioners found that inclusive design was within practical reach for them. The final evaluations showed acceptance of and enthusiasm for the visual design and the research methods offered.

FUTURE WORK

The next phase of this project is the departure from the prototype stage, starting the development of the application, incorporating everything learned from the prototype phase. Compared with existing toolboxes and resources, we believe the Include Toolbox is specifically suitable for SMEs in IT, due to its visual design, setup, and the difficulty level of the methods offered, more so than other toolboxes. However, this is something that needs to be further validated with a comparative study of toolkits amongst IT practitioners.

Although the first informal evaluations have been positive, more formal empirical study is needed to verify our claim that with the Include Toolbox, IT practitioners are able to perform user research with a quality adequate for their commercial purposes. Also we need to further substantiate the business rationale of inclusive design and user research.

To find out more about the need and use for background information and a physical book, further evaluations need to be done to find out if a book is the best form to present the information, and how much background information is in fact required by the SMEs.

ACKNOWLEDGMENTS

The current work was carried out as part of COMMIT, a Dutch national public-private research program on information and communication technology, project P02: Interaction for universal access (http://www.commit-nl.nl). We thank Bas Holleman for his help with the interaction design.

REFERENCES

1. Brand, R., & Rocchi, S. *Rethinking value in a changing landscape. A model for strategic reflection and business transformation*. Philips design (2011). http://www.design.philips.com/philips/shared/assets/design_assets/pdf/nvbD/april2011/paradigms.pdf
2. Coventry, L., Jones, E. "The role of tea parties to elicit technology requirements to support the mobility of older adults." *Proc. PETRA* (2012).
3. Cremers, A. H. M., Neerincx, M. A., & de Jong, J. G. M. "Inclusive design: bridging theory and practice." *Engineering Psychology and Cognitive Ergonomics, Applications and Services* (pp. 323-332). Springer Berlin Heidelberg (2013).
4. Curedale, R. Various titles, Design Community College Incorporated (2012-2013).
5. Dindler, C., Eriksson, E., Iversen, O. S., Lykke-Olesen, A., & Ludvigsen, M. "Mission from Mars." *Proc. IDC* (2005).
6. Dix, A. "Human–computer interaction: A stable discipline, a nascent science, and the growth of the long tail." *Interacting with Computers*, *22*(1), (2010), 13-27.
7. Eikhaug, O., Gheerawo, R., Plumbe, C., Berg, M. S., & Kunur, M. *Innovating with People- The Business of Inclusive Design*. Norsk Designråd (2010).
8. Gaver, B., Dunne, T., & Pacenti, E. "Design: cultural probes." *Interactions*, 6(1), (1999), 21-29.
9. Gillen, J., Cameron, C. A., Tapanya, S., Pinto, G., Hancock, R., Young, S., & Gamannossi, B. A. "A day in the life." *Early Child Development and Care* Vol. 177, Iss. 2, (2007).

10. Hanington, B. & Martin, B. *Universal Methods of Design: 100 Ways to Research Complex Problems, Develop Innovative Ideas, and Design Effective Solutions*. Rockport Publishers (2012).
11. http://designingwithpeople.org
12. http://filter.ucdtoolbox.com
13. http://www.55plustoolbox.nl/Toolbox.php
14. http://www.hcdconnect.org
15. http://www.inclusivedesigntoolkit.com
16. http://www2.gov.bc.ca/gov/topic.page?id=157D6555F8744850A439544F41727402
17. IDEO, *Method Cards: 51 Ways to Inspire Design*. Palo Alto (2003).
18. Jongerius, P., Offermans, A., Vanhoucke, A., Sanwikarja, P., & van Geel, J. *Get Agile! Scrum for UX, Design & Development*. BIS Publishers (2013).
19. Kayal, A., Brinkman, W.P., Zoon, H., Neerincx, M., Van Riemsdijk, M. B. “Using norms for social applications to provide value-sensitive support to families.” *Submitted for publication* (2014).
20. Langdon, P., & Thimbleby, H. “Inclusion and interaction: Designing interaction for inclusive populations.” *Interacting with Computers*, *22*(6), (2010). 439-448.
21. Markopoulos, P., Read, J. C., MacFarlane, S., & Hoysniemi, J. *Evaluating children's interactive products*. Morgan Kaufmann (2008).
22. Nielsen, J. *Thinking Aloud: The #1 Usability Tool*. NNgroup (2012). Via http://www.nngroup.com/articles/ thinking-aloud-the-1-usability-tool/, retrieved 2013.
23. Nielsen, J., Landauer, T.K. “A mathematical model of the finding of usability problems.” *Proc. CHI, INTERACT* (1993), p.206-213
24. Nielsen, L. *Constructing the user. Human-computer interaction: theory and practice*. Part II 2 (2003) p.430
25. Ozcelik-Buskermolen, D., & Terken, J. “Co-constructing stories.” *Proc. PDC* (2012).
26. Information about the prototype is available upon request from info@hannazoon.com.
27. Razak, F. H. A. *Single person study: Methodological issues*. (Doctoral dissertation) Lancaster University (2008).
28. Ries, E. *The Lean Startup: How today's entrepreneurs use continuous innovation to create radically successful businesses*. Random House Digital, Inc (2011).
29. Seffah, A. “Bridging the educational gap between SE and HCI: what software engineers should know?” *Proc. IET* (2004), 88-95.
30. Sluis-Theischeffer, W., Bekker, T., & Eggen, B. “Adding user creativity to the UX toolbox: Exploring the use of Creative UX methods.” *Proc. CHI Netherlands* (2009).
31. Spencer, D., *Card Sorting*, O'Reilly Media, Inc. (2009).
32. Tidd, J., & Bessant, J. *Managing innovation: integrating technological, market and organizational change*. Wiley (2011), Chapter 1.3.
33. Van Doorn, F., Stappers, P. J., & Gielen, M. “Design research by proxy.” *Proc. SIGCHI* (2013).
34. Wheelchair teen in record-breaking backflip, *Metro UK* (2008). http://metro.co.uk/2008/11/24/wheelchair-teen-in-record-breaking-backflip-175437/

Chi sparks 2014

Breaking Usability Rules to Enable Reflection

Arnold Jan Quanjer[1,2]
a.j.r.b.quanjer@hhs.nl
[1] Communication & Multimedia Design Group
Academy for ICT & Media
The Hague University of Applied Sciences
The Netherlands

Maarten H. Lamers[2]
m.h.lamers@liacs.leidenuniv.nl
[2] Media Technology Research Group
Leiden Institute of Advanced Computer Science
Leiden University
The Netherlands

ABSTRACT

We propose a different view on user interface design and argue that, when designing an interface to encourage reflection, usability rules should be intentionally broken. In our study one of the most influential lists of usability heuristics is evaluated and assessed for its applicability in the design of interfaces that emphasize reflection. We explore the impediments of these heuristics and rephrase some in a complementary or even contrary way that fits the design process of an interface that specifically encourages reflection.

Author Keywords

Interaction design; user interface design; reflective design; usability.

ACM Classification Keywords

H.5.3 Human-centered computing ~ Heuristic evaluations

INTRODUCTION

On first sight there is nothing unusual about the bicycle depicted in Figure 1. But for an unsuspecting cyclist trying to ride it, there is a surprise in store. When steering to the right the bicycle turns to the left and vice versa. It is an exceptional challenge to ride it; in fact it is nearly impossible. But the experience is intense and may induce new insights on cycling.

Whereas a normal bike can be operated without a need to think, this bike induces thought about a many things: about gravity, about balance, about how you stay upright on a normal bike in the first place. This example shows that making the familiar strange can evoke questions that otherwise would not be asked. Just a little modification can make something a challenge to operate but can encourage reflection. Encouraging reflection on a task, action or topic is the aim of reflective design, which includes the design of reflective user interfaces.

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

Figure 1: The challenging bicycle. Rental bicycle from Gekke Fietsen ("Crazy Bikes"), Belgium. Designer and photographer unknown.

In our study we propose a different view on user interface design and argue that to encourage reflection. We argue that the rigid application of usability standards can be an impediment when designing interfaces that should encourage reflection. By reviewing common usability heuristics and analyzing their propensity to stimulate reflection, we derive and propose 7 design recommendations for reflective interfaces.

USABILITY

Interface design methods are often oriented towards efficient satisfaction of short-term goals. From this perspective the main purpose of an interface is to facilitate flawless interaction between man and machine [5]. A user interface that is designed according to optimal usability standards can be appropriate in accomplishing practical tasks that require a short-term, quick resolution.

Usability is considered to be a precondition for a fulfilling user experience, but by itself proved to be a too limited view on HCI design: the affective side of a user experience is not taken into account [1]. We argue that usability can also hold impediments for designing an interface that specifically encourages reflection.

REFLECTION

In their paper on designing for reflection Lars Hallnäs and Johan Redström introduce the term "slow technology", a design strategy for technology aimed at reflection and moments of mental rest rather than efficiency in performance [3]. They consider time to be an important factor in designing for reflection. When approaching design from a usability perspective tools should be created that get the job done, quickly and efficiently. The span of time to get something done should be as short as possible. But time could also be perceived as something to dwell in, thereby encouraging reflection. Slowness in learning, understanding and presence gives people time to think and reflect. Hallnäs and Redström call this "time technology" that through its "slow" character emphasizes the presence of time and enables reflection.

They introduce slowness and time as key aspects of reflective designs. The absence of conscious time perception in usability-driven designs leaves fewer opportunities for reflection. Slowness opens up the user for a conscious perception of time that, in turn, enables reflection.

THE IMPEDIMENTS OF USABILITY

In this section we explore usability heuristics from a reflective perspective and rephrase some of them into design recommendations for reflective interfaces. Jakob Nielsen's "10 Usability Heuristics for User Interface Design" [6, 7] are commonly recognized heuristics for user interface design. These heuristics, conceived in 1995 from a pure usability point of view on HCI, proved to be quite persistent. At present they are still quoted and used as a starting point for elaborating more extensive heuristic lists.

When discussing guidelines for any kind of interaction design it is unavoidable to address heuristics on usability. Every interaction designer would probably agree that usability clears the way for a more profound and complete user experience. But is this also the case if usability is considered from a reflective perspective? In this section Nielsen's influential lists of usability heuristics is evaluated and assessed for its applicability in the design of interfaces that should encourage reflection. We explore the impediments of usability and rephrasing some usability heuristics in a complementary or even contrary way that fits the design of an interface that makes you think.

In Table 1 we explore each of Nielsen's 10 usability heuristics [7] with respect to their reflective quality.

Nielsen's heuristics	Reflective quality
Visibility of system status The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.	Rich feedback is an important feature of reflective interfaces [4, 10]. "Reasonable time" is probably meant to be "as quick as possible". In reflective interfaces slowness [3] and delay [4] are considered to be a key feature enabling reflection.
Match between system and the real world The system should speak the users' language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.	Although it is important not to alienate the user, reflective interfaces can deliberately distort the real world to create a different perspective [10]. Always following real-world conventions, in a way the user is accustomed to, can impede reflection because it does not allow different interpretations. Presenting real-world conventions in an unconventional context or seemingly illogical order raises questions and enables different interpretations [2, 8, 10].
User control and freedom Users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.	A reflective interface allows user control over the meaning-making process [10]. But making mistakes can be a very fruitful source for creativity. Chance and making mistakes is closely related to freedom. Therefore user control and freedom should be interpreted as conditions that encourage reflection. It is impossible to make a mistake in a reflective process because any outcome is valid. The system should not correct the user in a constraining way.
Consistency and standards Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.	In reflective interfaces the users are explicitly encouraged to wonder about the information they get presented. Inconsistency and ambiguity can encourage people to rethink the meaning of information presented to them as long it does not obstruct the flow of interaction [10]. Rigidly following standard platform conventions can impede the reflective process. Presenting the user with new, unexpected options can shake up entrenched views and make users receptive to new perspectives.
Error prevention Even better than good error messages is a careful design that prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.	A system should work like it is programmed but intentional errors, designed to change the perspective of the user, can encourage reflection.

Recognition rather than recall Minimize the user's memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.	Making an effort to recall something can intensify the learning-process and makes information easier to remember.[1] Figuring out a, not so obvious, interface can encourage reflection. But friction and indistinctness should not obstruct the interaction.
Flexibility and efficiency of use Accelerators unseen by the novice user may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.	A reflective interface should be flexible and adapt to the specific needs of a user. When efficiency is only intended to reduce the time span of an interaction to a minimum, it can also hinder reflection [3].
Aesthetic and minimalist design Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.	From a designers point of view irrelevancy should be avoided but the designer or system should in their turn never rigidly impose what is relevant for the user. Reflective design allows users to maintain control of and responsibility for the meaning-making process [10].
Help users recognize, diagnose, and recover from errors Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.	Utilize error messages to convey rich feedback on a user action. Suggesting a too confined solution may limit the space of possibilities that a user can explore to consider better options [4].
Help and documentation Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.	A reflective interface should be self-explaining. Extra documentation should be integrated in the interaction and make the reflective process more profound instead of just explaining it [10].

Table 1: Nielsen's usability heuristics [7] (left column) and an analysis of their reflective qualities (right column).

After evaluating and assessing Nielsen's heuristics for their applicability in the design of reflective interfaces the following general similarities and differences can be defined.

A heuristic evaluation can be executed from a usability or reflective perspective.

Guidelines for reflective design should be developed beyond the preconditions of usability.

Unintentionally disregarding usability rules can lead to frustration. *Intentionally* breaking usability rules, causing friction, delay and deliberate slowness, can be a reflection-booster.

DISCUSSION

After a closer look, Nielsen's heuristics indeed prove to be a set of usability-orientated principles; they contain no direct applicable design principles that encourage reflection. Nonetheless, where not conflicting with the intended reflective features of an interface, they can be treated as preconditions for reflective design: after all, a reflective interface may be usable.

Overall, however, Nielsen’s heuristics proved to be too constraining instead of giving space to reflection. By establishing these impediments and rephrasing some of them in a complementary, sometimes contrary way the following recommendations for reflective design can be suggested:

Provide rich feedback to show the space of possibilities.

Use distortion of the real world to create a different perspective.

Use slowness and delay to enable reflection.

Never impose what is relevant for the user and allow users control over the process of meaning making.

Don't correct the user in a constraining way; making mistakes can be a fruitful source for creativity and learning.

Don't follow standard conventions but dare to present the user with new, unexpected options.

Recall to remember. Making an effort to recall something can intensify the learning-process and make information easier to remember.

[1] Making an effort to retrieve information from your memory is an efficient learning strategy. Tests enhance later retention more than additional study of the material. This phenomenon is known as the “testing effect” [9].

CONCLUSION

In our study we make a contribution to the field of designing interfaces that specifically encourage reflection. After conducting an explorative research on commonly accepted usability rules, we derived and propose 7 design recommendations. We realize that some of our argumentation in table 1 is based on our professional insight on interaction design and is yet not grounded in previous research. In future studies these assumptions and final recommendation, mentioned in the discussion, should be validated for their propensity to encourage reflection in the user.

We conclude that since reflective interfaces should be designed to *make you think*, they should be designed beyond the preconditions that are suggested by a pure usability orientated design. To be flawless and fast can be an obstacle to a profound reflective experience. Deliberate slowness and friction can encourage a user to make an extra effort and *make him think*.

REFERENCES

1. Anderson, S. P. (2011). *Seductive interaction design: Creating playful, fun and effective user experiences*. New Riders.
2. Graham, D. (1974). *Present continuous pasts*. In project 74 kunst bleibt kunst, Keulen.
3. Hallnäs, L., & Redström, J. (2001). "Slow technology: Designing for reflection." *Personal and Ubiquitous Computing 5(3)*: 201-212.
4. Jones, B. B. K. (2012). *Reflective interfaces: Assisting teens with stressful situations online*. Doctoral dissertation, Massachusetts Institute of Technology.
5. Krug, S. (2005). *Don't make me think*. Pearson Education.
6. Nielsen, J. (1994). *Usability engineering*. Elsevier Science & Technology.
7. Nielsen Norman Group. *Ten Usability Heuristics for User Interface Design*. www.nngroup.com (online resource: http://www.nngroup.com/articles/ten-usability-heuristics/. Accessed Nov 15, 2013).
8. Paik, N. J. (1974). *TV-buddha*. Collection Stedelijk Museum Amsterdam.
9. Roediger, H. L., & Karpicke, J. D. (2006). "The power of testing memory: Basic research and implications for educational practice." *Perspectives on Psychological Science 1(3)*: 181-210.
10. Sengers, P., Boehner, K., David, S., & Kaye, J. J. (2005). "Reflective design." In: *Proc of the 4th Decennial Conf on Critical Computing: Between Sense and Sensibility*, ACM: pp. 49-58.

chi sparks 2014

Ambient and Pervasive Design: HCI Education from UbiComp to Creative Design

Geert de Haan
Molenveld 44
7391ZD Twello
geert.de.haan@upcmail.nl

ABSTRACT
This paper discusses the design of the Ambient and Pervasive Design course as a practice-oriented course to introduce students to the basic techniques for building Internet of Things applications. AmbiPerv derives from a more extensive theory-oriented course entitled DevThis and we discuss the differences between the two considering the transition of the Media Technology curriculum towards a focus on Human-Centred Creative Technology.

Author Keywords
Ubiquitous computing; ambient intelligence; pervasive computing; social media; sensory applications; education; curriculum development.

ACM Classification Keywords
K.3.2: Computer and Information Science Education. H.5.2. User Centered Design.

INTRODUCTION
Developments in the area of HCI are proceeding in rapid succession with the consequence for teaching HCI curricula, that factual knowledge, book knowledge and the know-how to use particular programming languages and tools are gradually becoming less important whereas the ability to keep track of developments in the application area, the ability to do research to support design, and the ability to know how to develop conceptual solutions and translate them into demonstrable products are becoming more important [6][7][14].

HUMAN CENTRED CREATIVE TECHNOLOGY
Media Technology is a BA curriculum of Rotterdam University of Applied Science which focuses on Human-Computer Interaction in area of Media Design with a technical but user-centred focus. Based on the vision that HCI is rapidly moving out of the personal computer and out into the (real) world at large utilizing open data, sensory information, location-based awareness and intelligent interfaces, a course, entitled DevThis (short for: Development in Media Technology) was designed to address these changes and familiarize students with the new developments. The course was also designed to implement, albeit independently of the vision, a number of ideas concerning education, research and development in Media Technology. DevThis attempts to put teaching, designing a product and doing research in one course (see: [5][6]).

In DevThis, students are introduced to a number of new developments in the area of HCI and Media Technology. They are taught about applied scientific research with application development on the basis of scientific literature and resulting in a concept poster, a product prototype and a short scientific paper. Finally, the course itself acts as a facility for research and development.

In the past two years, some of the ideas with which DevThis had been developed did become common in Media Technology but also a relatively new concept arose: HCI or media design as *Human Centred Creative Technology*. With *Human Centred Creative Technology*, an additional focal point is introduced: the creative application of technology beyond the proverbial 'textbook'.

Teaching creative design requires, among other, extend the design practice with techniques to support novel solutions, such as, creative research tools, exploratory design and co-design. To facilitate creative design, often special workshop forms like pressure cookers are used and design laboratories like an open data lab, a sensor lab and a fablab. Finally, the notion of design as creating a product should move to the cyclic process of concepting, demonstrating and prototyping.

To support the transition towards the creative application of media technology, a new course was designed, based on DevThis but earlier; in the second year of the curriculum: Ambient and Pervasive Design or AmbiPerv.

DEVELOP THIS
Development in Media Technology (DevThis) was originally designed as an 8 credit research and development module for students in their third or fourth between the internship period and the BA thesis project. It was designed to achieve two educational goals at once: to teach advanced students to do scientific research and introduce them to the technological, design and scientific developments in the area of Human-Computer Interaction, such as social media and co-creation, personalisation, context sensitivity and location-based services, agile development, co-design, and the use of living labs and emergent design practices.

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

DevThis was also intended as a vehicle to perform research with two goals: first, to teach how to do research by actually doing just that. DevThis students were asked to develop a social, mobile and context-sensitive system to increase social cohesion, with the results that knowledge was gathered about how to do this, see: [5][6]. Similarly, when the new sensor lab/fablab/open datalab was opened in 2011, students were asked to create intelligent sensory applications to help manage the laboratory building and its facilities and devices. In this case, the research projects were used to acquire knowledge about how to do this. Also, a number of demonstrators were build to show what one can do with such laboratories.

In addition to gathering knowledge, DevThis also attempted to make the results of education projects useful beyond education itself. To avoid wasting effort, creativity, knowledge and actual results, the second research purpose of DevThis is to facilitate the accumulation of design knowledge. According to Troxler and Wolf [13], the concept of accumulation of knowledge (or ideas, creativity ...) is common to fabrication labs aiming to serve as innovation ecologies. Likewise, DevThis students are requested to leave behind documentation and examples for future re-use and extension of their product designs.

Develop This setup

The basic design for DevThis was that the module consists on the one hand of theoretical instruction about ubiquitous computing, including learning to read, study and possibly write scientific papers, and on the other hand it consists of a practical component where students work in teams of 3 or 4 on a ubiquitous computing project in which they have to develop a demonstrator and whilst motivating all their design choices considering software architecture, the frameworks used, the development methodology chosen, etc. etc. The classical teaching takes place in a number of lectures about developments in the research area and about research methodology which are supported by regular student assignments; so much as possible, assignments are setup to allow the students to link the our teaching at the scientific- and technological levels to recognisable elements in their own 'everyday' level of their experience, which is very much oriented towards the developments on the market and in the media. Among the concepts that students choose are automated attendance lists, information adapted to situated displays, tagging of clothes for suitable combinations, etc.

In addition to classical teaching, arrangements have been made to have students teach each other by means of lecturing their classmates, and presenting their own specific areas of expertise in workshops. Following Pask [10] this mode of learning is referred to as 'teachback', assuming that learning is facilitated by actively explaining and working with learning material. These are some of the lecture topics:

- Arduino as a development platform
- the psychology of design
- use of digital maps
- face recognition
- user profiling with sensors
- object recognition algorithms

Finally, students learn by being actively engaged in team research projects, which prepares them for their final thesis project and for working life, and they have the opportunity to do and learn from doing their own research project.

On the basis of a literature study, student teams develop their own conceptual solution and present these as a research poster about halfway though the module. Next, making a substantiated choice of development platform (Arduino, Apple, Android, etc.), tools, frameworks and the design and development methods (co-creation, Scrum, XP, etc.), the teams develop a prototype as a proof of concept. Finally, the teams present and document their findings in a report for re-use in follow-up research projects, and they may receive a bonus mark for writing a short paper equivalent of their regular project report. These are examples of concept demonstrators:

- Wifi broadcasting @ site
- Building access control
- Mobile money with NFC
- Indoor climate control
- A Bluetooth remote for old TV's
- Ubiquitous game with sensor data

DevThis evaluation

DevThis modules have been successful in teaching students how to do research in combination with doing research in relation to the new developments in the field. Compared to a few years ago, students know much more about the tools, techniques and frameworks for developing mobile and ubiquitous computing applications and both, the knowledge of the research and the ability to use scientific resources and methods have significantly improved [5].

There is one major disadvantage in DevThis with respect to supporting the *Human-Centred Creative Technology* approach, common to many 'knowledge-centric' approaches to creativity and creative design: the notion that creativity should be preceded by expertise [3][4].

Evidence in the ICT and Media area seems to suggest that successful application of creativity does not require complete knowledge of the application area but rather that it requires just a basic understanding of the area in combination with accessible tools [3][4][8]. As such, in order to support the development of creativity within a regular HCI or media curriculum, courses might better not focus too much on knowledge but on resources and tools.

AMBIENT AND PERVASIVE DESIGN

The *ambient en pervasive design* (AmbiPerv) course is a 2 credit second-year course which exemplifies how a course like DevThis which demonstrates various aspects of the vision of Human-Centered Creative Technology, may be used to guide the design of other courses. Second year MT students have been introduced to e.g. user interface design, software engineering and (PHP) programming but they have not yet been introduced to more advanced topics like e.g. research methodology, design patterns or sensory interfaces.

In AmbiPerv, students use the Arduino toolkit ([1]; see: http://arduino.cc) to work with sensors and effectors in IoT applications, running on the connected pc or as stand-alone on the Arduino board.

The purpose of AmbiPerv is to let students get acquainted with the Internet of Things (IoT) within the areas of HCI and Media Technology. First, students are introduced with the concept of IoT using video material from Bassett et al. [2] and other examples. IoT is related to the 'Make' movement and a link is created with the Arduino development platform with the well-known presentation of Massimo Banzi at Ted [1] exemplified with building the 'Blink' application with an Arduino kit. Experimenting with the Arduino kits is supported with references to resources and Arduino Crash Courses (e.g. [9][14]). Some of the students who had already followed the first-year optional FabLab course [12] sometimes acted as student-teachers.

Subsequent lessons introduce students to the basic concepts of working with electronics and stepwise refine the blink-application to reading sensors, transferring sensor information to a pc over the serial port with PHP and Node.js (see: http://semu.github.io/noduino/) and presenting the results in a graph that may be accessible over the internet, using the 'processing' programming language (see: processing.org). Lessons are designed such that students are not merely taught about how to get things done but rather where to locate additional resources and examples to create their own solutions.

In addition to introducing the concept of the Internet of Things, further explanation deal with various subtopics in more depth such as a brief history of Ubiquitous Computing, the semantic web and linked data, and various new approaches to designing media applications, such as co-creation and co-design [11], agile design, and IoT-specific design notions such as mashups and exploratory design [7].

In parallel to introducing Arduino and the Internet of Things, students do three assignments of increasing difficulty. Assignments start with copying or rebuilding the Arduino applications which deal with the basic techniques for building IoT applications in general:

- connecting the development board
- reading sensors
- writing effectors
- transferring data
- presenting information elsewhere

Next, students are requested to develop their own concept idea for building an IoT/Arduino application to sense whether the atmosphere is healthy, guided by a worked out example (see: http://learn.adafruit.com/tmp36-temperature-sensor/overview). The concept must utilise multiple sensors and use social media to influence the behaviour of people.

To develop a concept it is necessary to do both theoretical research, to establish what to measure for what purpose, as well as technical research, to establish how to measure and implement the concept.

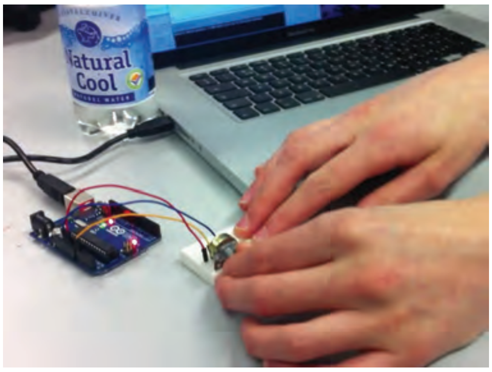

Figure 1. Cool Arduino project

Finally, students are requested to actually build a prototype as a 'demonstrator' of the feasibility of the concept, and make a videotape which is handed-in for evaluation.

AmbiPerv versus DevThis

In the ambient and pervasive design course, three aspects of the DevThis approach were skipped to address the fact that AmbiPerv students are less advanced in their studies then the DevThis students and because an introduction to the basic (programming) techniques for building IoT applications in AmbiPerv replaces the research goals of DevThis.

First, there is no need to provide a more or less complete overview of the new developments in HCI, such as agile and co-design methods, intelligent interfaces and metadata. Instead, rather 'tangible' examples are presented in an easy to grasp manner, like YouTube films.

Secondly, students are not requested to create a design based on extensive theoretical research; they are not required to read scientific papers, to argue why a particular solution has been chosen, and neither is there a requirement to scientific paper alongside the end-result in the form of a prototype application.

Thirdly, it is not required to follow and evaluate a particular design method and consequently, it is not possible to investigate the utility of various design approaches. Instead, students are led through a development process in which, in a step by step approach, all the essential or basic techniques to create an Internet of Things (IoT) application are presented, explained and evaluated. Student may focus on building without being bothered with scientific research.

DISCUSSION

The design of Ambient and Pervasive Design differs in various aspects from its origin Develop This. These changes have been introduced in order to meet the demands of presenting an IoT course in an earlier year in the Media Technology curriculum in combination with the wish to focus on solely proving students with the technical necessities to start building Internet of Things application rather then provide a scientifically more-or-less complete overview of the area in combination with opportunities for applied research.

Research opportunities are lacking in the current development of the course. In subsequent versions, it may be possible to introduce some form of research. Hopefully, when students are technologically better equipped, more interesting investigations may be possible, later on.

The evaluation and performance data about AmbiPerv is incomplete and, as such, only conclusions may be drawn from most of the actual lessons and from personal impressions.

Two conclusions are prominent: first, second year students enjoy building things with their hands as opposed to studying research papers. In this respect, we may conclude that a practise-based approach has a noticeable and very positive influence on student-motivation. This conclusion supports the notion that education should support rather than hamper creative exploration and experimentation as suggested by [3][4] about learning in children and by [8] regarding design students. Even thought second year AmbiPerv classes are more crowded then the third year DevThis classes, it turned out both easier and more pleasant to teach AmbiPerv.

A second effect of AmbiPerv is that students become familiar with notions such as sensory systems, intelligent interfaces and mashups much earlier in their studies which creates better opportunities that students are eventually going to design creative and innovative applications. It is remarkable how much these students differ from the DevThis students who attempted to develop Arduino applications, a few years ago, without any teaching support. AmbiPerv shows that a tightly structured course that takes students by the hand makes it possible to teach the technical skills for creative design early on in the curriculum. It remains to be seen, however, if the increased skill levels in an early stage will eventually compensate for a diminished level of abstraction and less knowledge, later on. In similar vain, one may ask what will remain of an advanced course like DevThis when it is stripped from advanced topics like research methodology, design methods and, with AmbiPerv, interfacing with sensors and effectors is put in a separate course.

Develop This and Ambient and Pervasive Design differ in their position in the curriculum but also on ideas about how creative design should be encouraged in design courses, such as: based on minimal examples or a complete overview, self-guided scientific exploration or guided by assignments, directed at research or at basic how-to-do-it skills to stimulate creativity.

Overall, it seems safe to state that presenting students with the basic techniques earlier-on in their studies will have a positive effect on the levels of creativity in the design capabilities in future students, supporting the transition from knowledge-based design to media design as *Human Centred Creative Technology.*

ACKNOWLEDGEMENT

I am indebted to my former colleague Rimmert Zelle for co-creating (or co-designing!?) Ambient and Pervasive Design as well as its predecessor UbiComp in the Mobile Life minor.

REFERENCES

1. Banzi, M. (2012). How Arduino is Open Source Imagination. *Ted Global presentation*, 26 June 2012, Edinburgh.
2. Bassett, T., Casden, A., Fitzloff, S., Jain, A. and Michel, C. (2012). *Connecting*. Video documentary.
3. Brennan, K. (2013). “Computational Creativity: How New Tools and New Communities Support Young People as Designers of Interactive Media.” *Keynote at CSERC2013, the 3rd Computer Science Education Research Conference*, 4-5 April 2013, Arnhem, The Netherlands.
4. Giannakos, M.N., Jaccheri, L. and Proto, R. (2013). “Teaching Computer Science to Young Children through Creativity: Lessons Learned from the Case of Norway.” In: *Proc. Computer Science Education Research Conference - CSERC '13*, 4-5 april 2013, Arnhem, The Netherlands.
5. de Haan, G. (2011). “DevThis: HCI Education beyond Usability Evaluation.” *Proc. Chi Sparks 2011*, 23 juni 2011, Arnhem.
6. de Haan, G. (2012) “DevThis - Education in the age of the Internet of Things.” *Proc. EADiM2012*, 22-24 November 2012, Graz, Austria.
7. de Haan, G. (2013). “A Vision of the Future of Media Technology Design Education - design and education from HCI to UbiComp.” In: *Proc. Computer Science Education Research Conference - CSERC 2013*, 4-5 april 2013, Arnhem, pp. 66-72.
8. van der Helm, A., Stappers, P.J., Keyson, D., & Hekkert, P. (2010). “The ID-StudioLab 2005-2010: Further developing a creative research environment.” In: *Proc. Design Research in the Netherlands*. 20-21 May 2010, Eindhoven, The Netherlands, pp 65-78.
9. Markaban, F. (2011). *Crash Course Arduino. DevThis student lecture*.
10. Pask, G. (1975). *Conversation, Cognition, and Learning*. New York: Elsevier.
11. Sanders, E.B.-N., and Stappers, P.J. (2008). “Co-creation and the new landscapes of design.” *CoDesign*, 4 (1), 5-18.
12. Troxler, P. and Mostert-van der Sar, M. (2012). *FabLab - Wat heb je nodig om (Bijna) alles te kunnen maken. Modulerwijzer FabLab*. Instituut CMI. Hogeschool Rotterdam.
13. Troxler, P. and Wolf, P. (2010). “Bending the Rules. The Fab Lab Innovation Ecology.” *Proc. of the the 11th International CINet Conference*, Zurich, Switzerland, 5-7 September 2010.
14. van Turnhout, K., Leer, S., Ruis, E., Zaad, L. and Bakker, R. (2012). “UX in the Wild: Experience Blend & Embedded Media Design.” *Proc. The Web & Beyond 2012*, 26 Sept. 2012, Amsterdam.

Building UX capabilities at a technology firm: A shared success story

Jesse Grimes
Informaat
P.O. Box 789
3740 AT Baarn (NL)
jesse.grimes@informaat.nl

Anneke van de Langkruis
Priva
P.O. Box 18
2678 LC De Lier (NL)
anneke.van.de.langkruis@priva.nl

ABSTRACT

Over the past several years, experience design consultancy Informaat and climate control specialist Priva have delivered a series of successful projects. Informaat's UX design, service design and recruitment and placement services have supported the growth of what is today a mature, in-house design capability. Our learnings – along with those of our Priva – offer other UX practitioners insights that we hope will deliver value to their roles.

Author Keywords

User experience; experience design; design management; product design; service design.

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION

In 2008, Informaat began what would become a rewarding client relationship that continues to this day. What began with a UI design project grew into a service design one, and since then we have supported Priva in creating their in-house UX design capabilities, through further tactical projects and on to staff selection and recruitment. In this shared paper, we present our observations and learnings as a design agency supporting Priva's UX maturation, and our Priva's learnings throughout the process. We hope to share our combined experience earned through more than 20 completed projects with the wider community of UX practitioners.

PRIVA AND INFORMAAT

Priva is a long-established Dutch company which supplies systems (hardware and software) to manage indoor environments, either for residential or office buildings or horticulture. Informaat are a customer experience design consultancy that started in 1986, and carries out service design, interaction design, and content design projects.

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

THE INITIAL BRIEF

In late 2008, we were briefed by Priva to address three specific requirements: Create (user interface) consistency between their different products, improve their usability and learnability, and support their localization and translation (allowing them to distribute products internationally and in different languages). However during our initial project — which was dedicated to the UI of a specific product — it became apparent that any user experience design efforts would need to be broader in scope. We proposed a service design project to deliver broader, more informed insight on the customers and end users, rather than immediately focusing on interface design issues. This involved an in-depth look at all the customer and end-user touchpoints with Priva products and services, and resulted in the creation of a service ecology.

LAYING THE GROUNDWORK WITH SERVICE DESIGN

Together, we conducted the service design project, and achieved a better understanding of the scope of their services, the way products were used, and the interactions between different roles in the service lifecycle. These were visualized in the form of personas, customer journeys, and service ecosystems.

Led by a newly-hired internal brand manager within Priva, this project represented the introduction of a truly new way of thinking for the company. For the first time, there was a shared understanding - across the organization - of who their customers were, supported in large part through a series of service design workshops. Customer insight was no longer the exclusive territory of sales and marketing, and a series of successful service design workshops disseminated this new form of "design thinking" across the organization.

APPLYING STRATEGIC DESIGN TO TACTICAL PROJECTS

In the years following our first service design engagement, the focus shifted towards UI design projects on specific product lines. These coincided with efforts to relaunch these products, and built upon our service design work. One particular project successfully demonstrated how UX design delivered excellent results.

"Touchpoint" is a small, touch-screen device that allows installers and building operations managers to directly configure aspects of a building's environmental control systems, without using full-screen, computer controls.

During this three-month project, design considerations were put at the core of the project; rather than squeezing design activities into an IT-oriented Rational Unified Process (RUP) or waterfall approach, adequate time was allowed for user research, and to test and iterate designs. For example, in situ user research carried out in the Netherlands, the UK and Germany, helped to precisely pinpoint the exact users and precise use cases in which the product would be used. This allowed the team to significantly narrow the scope of requirements to deliver a precisely targeted, fit-for-purpose product that didn't overlap with other products. In turn, development activities could be focused on building only what was necessary, saving considerable cost and time.

Usage scenarios developed in this project built upon what was learned during service design work, and were further honed by user research. In a further example of how this insight improved the UX of the end-product, research uncovered that end-users sometimes carried out configuration or troubleshooting activities on the device while on the telephone with a third party. This meant that the interface and navigational structure needed to be simple enough to support voice-based guidance. As with many aspects of this project, this UX improvement would not have made it into the final product if design didn't play its new, significant role within the organization.

A NEW DESIGN APPROACH PAYS OFF

The service design project was considered a great success by the design team and high-level management, but the wider recognition of the value of this new design approach came when stakeholders throughout the business started to see very positive reaction to the new UI designs that were developed and released to customers as part of our later projects. For them, the activities that generated things such as personas and service ecologies could be seen to have delivered a tangible, bottom-line result.

EMBEDDING UX CAPABILITIES

Since our initial engagement, we have delivered more than twenty projects for Priva, bringing UX expertise to their organization and supporting the team's own growth and development. Between 2008 and 2012, this included work on their corporate website, graphical asset library, software application identity, web applications and product innovation activities.

Over time, however, it was Priva's goal to become self-sufficient. There was the recognition that too much UX knowledge and specific experience was being carried by external staff, and risked being lost when projects were complete. Moreover, UX had - in the meantime - been chosen to become an internal core competency of the client.

In response to these demands, we adapted our engagement to assist them in building their in-house UX capability, both in terms of recruiting staff as well as providing concrete guidance on how to put it into place. This aligns with Informaat's approach to our client engagements as well, which we describe as "deliver, co-create and support" (in Dutch: "voordoen, meedoen, zelf doen"). In other words, we often carry out work on our own for clients with no in-house UX capabilities, however we progress to shared projects and finally aim to leave clients with the tools, skills and expertise to run projects independently.

Another important role we played was to raise the profile of the design team as it grew and took on greater responsibility. In our role as external consultants carrying out long-term projects, we built relationships with internal stakeholders, and continually reinforced the role and value of design.

In Priva's case, our service design and UI design projects culminated in efforts to find and recruit staff to build their internal UX design department. We carried out these activities over the course of 2012, and currently, half of the six-person team was placed by ourselves. More and more, we find that this approach is desired by our clients. The capability to execute UX projects independently, or with the support of an agency, is preferable to long-term reliance on a third party.

THE VALUE OF UX TODAY

Priva's relatively recent focus on UX fits within a broader historical context. Since the late 1950s, they have recognized the value of good design, and have maintained a relationship with TU Delft, centered on product design. In fact, design has reaped business benefits too; since that time, more than €20m in business value can be attributed to this focus on the value of good design. Today, UX is considered internally as a key differentiator and USP for Priva, and the company aims to continue improving UX to set themselves apart from their competitors.

INFORMAAT'S LEARNINGS

Our on-going relationship with Priva has taught us several things that we feel are worthwhile sharing with practitioners such as ourselves:

- Consensus, high-level design representation, and patience are all required to introduce design thinking throughout an organization. And it can only be accomplished with someone on the client-side leading the effort.
- A "deliver, co-create and support" approach meets a client's short- and long-term goals, yet still offers the possibility of fruitful, multi-year relationships which can include activities such as organizational consultancy services, and the placement of contract-based designers on an interim basis.
- Unlike consumer-focused industry sectors such as travel, retail and finance, B2B technology and industrial clients are often lower on the UX maturity scale, and offer greater opportunities for varied types of UX work.
- A workshop-based approach (especially during service design and strategy-level work) allows multidisciplinary teams to visualize and co-create in a creative and unstructured environment. The results (which engage and surprise participants in equal

measure) can be shared and hung as large-format visuals, reinforcing the work long after the workshops finish.
- The ability to follow up a strategic project (service design) with tangible, concrete ones (UI design) significantly raised both the impact and long-term durability of our work. Too often, standalone strategy projects that don't result in actionable deliverables fail to deliver long-term value.
- Those wishing to find UX work (especially recent design graduates) can find promising opportunities for challenging, rewarding in-house design work away from the traditional marketing- and communication-led agency world.

PRIVA'S LEARNINGS

- Service design is an ideal way to introduce design thinking and the value of UX within an organization, because it can be carried out at a strategic level, and see its outcomes applied tactically in the future.
- A documented UX design process is crucial for ensuring that design plays its role in projects at the right time, every time. In addition, explicitly including design activities in projects by making it part of a structured process helps reduce the risk that it is ignored or overlooked.
- Design deliverables that can be broadly applied across different stages of different projects (such as personas) serve to reinforce the value of UX consistently. Training the audiences of these deliverables - such as developers - pays dividends too.

CONCLUSION

Looking back on many years of fruitful collaboration between Informaat and Priva, we have seen firsthand the value that UX awareness and efforts can deliver to an organization. Our approach - from strategic to tactical, and from fully outsourced to in-house - has paid dividends in terms of the profile and responsibility of the internal design team. It has also delivered learnings that can be applied by other practitioners seeking long-term client engagements, and for internal designers seeking advice on how to position UX within their organization's overall strategy.

Chi sparks 2014

Concept Plan for User Testing in Scrum Projects

Sirvan Ugur
Interaction Designer at Acato
Veenkade 48, 2513EH
The Hague, the Netherlands
sirvan@acato.nl

Swie Oei
Scrum Master at Acato
Veenkade 48, 2513EH
The Hague, the Netherlands
swie@acato.nl

Arjen de Bruijn
Interaction Designer at Acato
Veenkade 48, 2513EH
The Hague, the Netherlands
arjen@acato.nl

ABSTRACT

In this paper, Acato (design agency for interactive media, based in The Hague, ed.) describes a concept plan for user testing in a scrum project. Scrum is a design method in which a multidisciplinary team consisting of designers, developers, a Scrum Master and a Product Owner, design and develop features and functionalities during iterative sprints (i.e., repeated work cycles). The guidelines for user testing by Belk [1] during scrum are:

- High involvement of the project team
- Minimal documentation
- Lean sessions to set up a research plan and analysis

With these guidelines in mind, a concept plan has been drawn up for user testing in scrum. Using the plan, the team analyzes the results based on the research goals identified in it. The tests are kept lean and are ideally observed collaboratively. The points of action from the results are used to plan the required changes for the next sprints. Integrating user testing in a scrum project allows usability issues to be addressed at an early stage in a development, which improves the product gradually throughout the development process.

Author's Keywords

User research; Scrum; User testing; Interaction design; Development; Lean; UX; Agile.

ACM Classification Keywords

H.5.2 User Interface: Evaluation/methodology

INTRODUCTION

Acato has been successfully conducting user tests for over 10 years. With the introduction of new design methods such as scrum, Acato has identified a need for a new, leaner method of user testing. This paper describes a recently discovered approach that Acato encountered during the UX People event in London, 1 November 2013. At the event Nina Belk, User Research Manager at LBi delivered a workshop entitled 'Agile-friendly user testing' [1]. As Acato had already carried out several scrum projects the workshop attracted our attention. We were able to acquire valuable skills into conducting agile user research, which in turn can also be used well in scrum projects. This report describes how Acato set up a concept plan inspired by Belk's workshop and drew up guidelines based on her instructions for an agile friendly approach.

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

USER TESTING

Acato has been conducting qualitative user tests [2] since 10 years. This has given us important insight into user experience and user behaviour. Depending on the time and budget available, and in light of the significant advantages to be gained, Acato clients are advised to test designs. This can be done either extensively or pragmatically (e.g. guerilla testing [3,4]). During the past 10 years we have developed a toolkit for user testing internally comprising a research plan, an observation template, a useful setup and a standard routine for recruiting participants.

Publications such as Steve Krug's 'Rocket Surgery made Easy' [5] have become standard reading for Acato employees and several different methods and tips have been implemented. Acato has based some of the ingredients of its own toolkit not only on Krug's methodology, but also on articles to be found in Nielsen Norman Group [6].

The most important components of user tests at Acato are:

- A mock-up of the website or application to be tested
- An observer and an interviewer
- A test script
- A number of scenarios and tasks for the participants to complete during the test
- Card sorting
- Pre-test and post-test questions

Within Acato user testing has the following 7 phases:

1. Research goals & research plan

Approximately 3 weeks before testing: determining the research goal and the research questions. The outcome is a detailed research plan to be reviewed by the client.

2. Recruiting participants

Approximately 3 weeks before testing: recruiting participants representing future users of the product. This is done by either the client or Acato. Shortly before the actual test a reminder is sent to the participants to increase their sense of involvement and confirm their attendance.

3. Pilot test

Approximately 1 week before testing: conducting a pilot test to optimize the setup (mock-up) and the test script.

4. Preparing the venue
Approximately 1 day before testing: setting up the location, testing the camera and recording/screen sharing software.

5. Interviews
The actual tests: the advice is to test 5 to 10 participants [2]. Each interview takes 45 minutes to an hour. User tests can take up to a total of 3 days.

The tests usually consist of:

- An introduction, during which the interviewer introduces him/herself and the observer, and briefly explains the research goal;
- A number of pre-test questions, to break the ice and gain insight into any relevant experience the participants may have;
- A first impression of the look & feel, followed by card sorting;
- A number of scenarios that put the participant in the right frame of mind, combined with a few tasks [7] that implicitly contain the research questions;
- Final post-test questions, often used to find out what seemed most and least useful to the participant, and to get their overall opinion.

6. Results
Approximately 1 week after tests: the analysis of the results. The outcome is a detailed report including recommendations to the client.

7. Re-design
Approximately 2 weeks after testing: making changes to the design, based on the recommendations.

Conducting user tests takes quite a long time and the overall process is intensive. During scrum projects, time is often limited. This led us to search for leaner user testing methods.

ABOUT SCRUM

Scrum has been developed over the past 20 years from a project management method solely for software development, to a method that also integrates design in the process. This evolution has led to the creation of a multidisciplinary team, with designers, developers, a Scrum Master (SM) and a Product Owner (PO). The PO is usually a representative from the client and supported by other stakeholders. Together, they safeguard the end goal of the project. This process and method are explained in 'Get Agile' [8], by Pieter Jongerius, a publication widely considered a useful manual for scrum.

Scrum consists of a sprint 0, followed by multiple further sprints. In each sprint, features and functionalities are designed and developed. After each sprint, iterations on the design can be made, as new insights are acquired.

Sprint 0

The first sprint is essential for the subsequent sprints. In this sprint, preparations are made, the scope is determined and the number of sprints, sprint weeks and sprint days within a sprint are laid down. In fact, sprint 0 forms the framework for the rest of the scrum project.

As scrum is all about eliminating waste and focusing on the creation of a Minimum Viable Product or MVP (formally defined, for example, as: a product that has only those features (and no more) that allow you to ship a product that resonates with early adopters; some of whom will pay you money or give you feedback [8,9]). The Product Backlog is a prioritized collection of known features, functionalities and wishes. In addition to MVP as a design goal, Acato also strives to design Minimum Lovable Products (MLP) [10, 11]: products people love to use.

With scrum, the end user is placed at the heart of the process. All features, functionalities and wishes are therefore documented as user stories: short descriptions of a feature describing the added value for the user. If certain stories do not have an added value, the team may reconsider if a given feature should be in the included in the Product Backlog or not.

If the client wants to include user testing, preparations for this should be made in sprint 0. The actual goal of the tests is determined during the sprints, according to the most important features (as, in most cases, this is what you want to test).

Sprints

In the first day of the sprint, the team defines the sprint with the most important user stories from the Product Backlog. These stories form the basis for the deliverables of that sprint.

The end of each sprint is characterized by a demo, when the team presents the deliverables: a working version of a part of the product. Stakeholders are invited to see the result and give their feedback. Any new insights or adjustments can then form new input for the Product Backlog.

Once all sprints have been completed, the end result is presented to the stakeholders, and the product is delivered.

In short, scrum is characterized by the following benefits:

- Eliminating waste by focusing on the most important features (MVP) facilitates a shorter time to market
- A dedicated team increases focus
- Flexibility: as the PO controls the stories in a sprint
- Transparency: the whole team has insight in a sprint's progress
- The PO as part of the team facilitates direct communication between the scrum team and the client

INTEGRATING USER TESTING IN SCRUM

Prior to conducting user testing in a scrum project, there are a number of issues that first need to be addressed:

Identifying the research goals and analyzing the results

First of all, during a scrum project, there is insufficient time to cover the phases described above. The first phase, setting up and finalizing the research plan can take up to 3 weeks.

This phase would cover an entire sprint. It is therefore important that this phase is kept as short and as lean as possible. In her workshop, Nina Belk illustrates a creative session with the project team to draw up the research plan. Not only can this session be held in one or two hours max. (after which, ideally, the research plan merely needs to be perfected), but it also increases the involvement of project members, which is one of the important characteristics of the scrum method.

This approach can also be used for analyzing the results (after user testing): instead of having the results analyzed by the agency and reviewed by the client, they are analyzed together with the client and the project team (or part of the team). Ideally, the only thing remaining after such a session is to summarize and organize the results, and make sure everyone receives a copy.

It will probably not always be efficient to carry out these sessions with the entire team: to save time and to keep the process lean, the analysis can be done by the SM and PO alone.

Interviews and observation

The tests are normally led by the interviewer. The observer takes notes and may ask additional questions. At Acato, some user tests have been conducted while the client watched the interviews in another room. This has proven to be a useful way of demonstrating the value of user testing to the client, as they can actually see users encountering problems, or giving positive feedback. During scrum, other team members, for example the developers, may be also asked to observe. It is valuable to not only involve the client and team members in the observation, but to also let them keep track of the findings.

However, we appreciate that the involvement of the client and team members to this extent may not always be practically possible. Moreover, there is the risk that the perception of people closely involved in the realization of the product may be biased.

Documentation

Finally, there is efficiency to be won when it comes to documentation throughout the project. The level of detail in documentation varies according to the client and the context: some circumstances necessitate a clear report, due to the number of people involved in a project; other times, an email with summaries and recommendations may suffice. In scrum user testing, it is important to keep documentation to as little as possible: the research questions, together with the most important findings and recommendations will suffice. The recommendations form new (or adjusted) stories that can then be added to the product backlog, which in turn can then be added to a new sprint.

Guidelines

To recapitulate: to ensure user testing can be applied successfully in a scrum project, it is important to bear in mind the following guidelines based on Belk's instructions [1]:

- The project team should be actively involved, not only during the sessions but also during the drawing up of the research plan and the analysis
- Sessions should be lean and short, focusing solely on those elements requiring attention
- Documentation throughout the testing process should be kept to an absolute minimum

CONCEPT PLAN

Using the guidelines outlined above, the following plan was drawn up for conducting user testing in future scrum projects.

Research goals workshop

As already stated, in a scrum approach, the setup and processing of the research questions is carried out by (at least) the PO and the SM. A brainstorm workshop is held to draw up the research questions; covering elements of the design that are unconventional or uncertain, and hence in need of testing w.r.t. the user experience. In addition to user experience, commercial considerations may also be important.

Some examples of research questions are:

- How is the visual design of the site perceived by the target group?
- How is the navigation perceived?
- What problems do users encounter when using the input form?

The easiest way to identify questions during the workshop is to print and display print screens of the elements/pages to be tested for the team. The designers go through the screens while project team members brainstorm for relevant research questions. These questions are written on sticky notes, together with the page number. Filtering and categorizing these sticky notes leads to a collection of research questions, which can subsequently be processed into scenarios and tasks [7]. These scenarios are pitched to participants, who are asked to complete a task. The tasks should implicitly contain the research questions, while the scenario puts the participant in the right frame of mind. The outcome of this session is the test script.

Recruiting participants

Participant recruitment depends, for example, on the available time and budget. For each sprint, the advice is to test 5 participants. With less than 5, there is a chance that the results will not be representative of the target group, while testing more than 5 decreases the return on the time invested compared to new insights gained, [2]. If the results and findings diverge, stand-by participants may be called upon.

Interviews

During the interviews the interviewer guides the participant through the script. The interviewer's attitude should be sympathetic, skilled and humble, so that the participant feels comfortable and not compelled to achieve. By sticking closely to the script, the interviewer avoids giving away too much information, or steering the test in the wrong direction.

It is also important to ask open-ended questions, to establish the participant's initial reactions and findings [12].

Location

A dedicated test room close to the observation room is preferred. This enables you to ensure the setup is comfortable and productive for both interviewer and participant. It may, however, at times be difficult finding participants willing to come to a given location. In this case, it is possible to interview participants in their own workplace, or even at home.

Observation workshop

During the interviews, team members (including stakeholders) are situated in another room, with the screen prints being tested, displayed on a wall. Each observer is asked to note any findings on sticky notes, including the name of the screen (S) and the number of the participant (P) (e.g. S1P1). It is important that team members not only note the issues they identify, but also to try to interpret what is going on and why.

The outcome of each session is a number of sticky notes with findings noted (per screen).

User testing at the participant's own home or workplace obviously makes observation by the team more difficult. Screen sharing software (e.g. Skype) enables team members to observe and take notes remotely. Where this is not possible, an observer should be present.

Analysis workshop

After conducting 5 user tests, the team comes together for a debriefing workshop to analyze the findings. Instead of drawing up a detailed report with all the findings (as is done in conventional user testing), a brief email with a clear and workable action plan based on the most important findings is sent to the project team.

Making changes to the design

Using the action points from the user tests, the required changes are written as user stories and added to the product backlog. These stories may then be included in the plan for the next sprint.

DISCUSSION

This paper outlines a concept plan for user testing in a scrum project. The hypothesis is that this can be implemented in scrum projects. The involvement of the entire scrum team is an essential characteristic of the plan, and beneficial to all disciplines in the team. It increases the overall understanding during projects, and most probably the efficiency of the problem solving process.

However, a few implications of this high level of involvement are to be expected. The following points should be taken into account:

- Scrum was initially an agile programming methodology. Several articles have been written about the implications of involving the UX team in scrum in general [13]. The preparations for user testing and other UX activities should be well planned and integrated in the scrum process.
- Clear team briefing is essential for each specific task, e.g. taking notes during observation, including not only the 'what' but also the 'format' of the task and 'why' it is important, as well as 'how' the result will affect the design or process.
- Careful selection of team members for each workshop is important. Firstly, the intensity and workload can be high during a scrum project, making involvement of all members sometimes impossible. Secondly, every discipline and team member has their own specialist field which may not always be necessary or useful during a particular workshop.

Throughout the process it is important to consistently evaluate the benefits of high involvement versus keeping it lean.

Minimizing documentation may also be difficult. Most designers spend considerable time documenting. They are paid for doing it and clients often want documentation deliverables [14].

Overall, integrating user testing in a scrum project allows usability issues to be addressed at an early stage in the development, whereby the product gradually improves throughout the process.

The hypothesis is that user testing as described above will add significant value to scrum projects and Acato looks forward to applying this new way of user testing on projects in the near future.

REFERENCES

1. Belk, N. *Agile-Friendly User Testing*. www.slideshare.net/NinaBelk/agilefriendly-user-research-nina-belk-ux-people-2013
2. Nielsen, J. *How Many Test Users in a Usability Study?* www.nngroup.com/articles/how-many-test-users
3. Simon, D.P. *The Art of Guerilla Usability Testing.* www.uxbooth.com/articles/the-art-of-usability-testing.
4. GOV.UK. *Guerilla testing: Getting input into products and services.* www.gov.uk/service-manual/user-centered-design/user-research/guerilla-testing
5. Krug, S. *Rocket Surgery Made Easy: The Do-It Yourself Guide to Finding and Fixing Usability Problems.* Pearson Education, US, 2009.
6. Nielsen Norman Group, Evidence-Based User Experience Research, Training, and Consulting. www.nngroup.com
7. McCloskey, M. *Turn User Goals into Task Scenario's for Usability Testing*. www.nngroup.com/articles/task-scenarios-usability-testing
8. Jongerius, P. *Get Agile*. BIS Publishers, NL, 2012
9. Ries, E. *Minimum Viable Product: a guide*. www.startuplessonslearned.com/2009/08/minimum-viable-product-guide

10. Austin, J. *Workshop From Minimum Viable Products to Minimum Lovable Products: finding your heart*.
11. Walter, A. *Designing for Emotion*. A Book Apart, US, 2011
12. Pernice, K. *Talking with Participants During a Usability Test*. www.nngroup.com/articles/talking-to-users
13. Budwig, M., Jeong, S., Kelkar, K. "When user experience met agile: a case study." *CHI EA '09* (2009), 3075-3084
14. Gothelf, J. *Lean UX: Getting Out Of The Deliverables Business.* uxdesign.smashingmagazine.com/2011/03/07/lean-ux-getting-out-of-the-deliverables-business

DEMOS & POSTERS

chi sparks 2014

Giving user-generated content back to the users: Testing data de-centralisation with active content creators

Amy Guy
School of Informatics
University of Edinburgh
amy.guy@ed.ac.uk

ABSTRACT
Our poster gives an overview of on-going work which is part of a broader effort to enhance and facilitate collaborative opportunities for digital content creators. We describe *creative media production social machines,* which gives a perspective on the user-generated content ecosystem that changes how we consider content creators and the technical systems they use. We observe individual content creator behaviour within social machines - understanding their behaviour from the bottom-up rather than top-down - in order to inform better design of systems which support their interactions. We illustrate on-going work in formalising certain observations, to provide a basis for designing and developing systems which can successfully co-operate in the creative media production social machine space.

Author Keywords
Social machines; user-generated content; metadata; creative media; user-centric design; identity; online communities.

ACM Classification Keywords
H.1.2. Information Systems: User/Machine Systems.

INTRODUCTION
Within the online user-generated content ecosystem, there are many *active* content creators. For them - amateur musicians, independent film-makers, artists, storytellers - content creation is a core part of their lives and perhaps a means to an income. They are actively seeking new audience members, and are likely to understand the ones they've got fairly well. They collaborate with other creators, work to improve their output, and cultivate one or more online personas through which they present their content.

They do not confine themselves to one website or community, but spread versions of themselves across several. Their success depends on both social and technical factors; how they present themselves, how their audience receives their content, how the platform(s) on which they publish permit their content to be consumed, shared and re-used, and how different content-host and social networking platforms may interact with each other.

We therefore take the perspective that these content creators are participants in one or more *social machines*: critical *parts* of the systems in which they operate, rather than merely users. Their desires and needs transcend individual websites, applications or communities, and we argue that this should be taken into account when improving existing systems or designing new ones.

Social Machines
Social machines are systems for which human and computational components are complementary and equally critical: "processes in which the people do the creative work and the machine does the administration" [1]. They exist already on many different scales and in many different domains, often evolving, responding to technological and social developments, and interacting with each other [8, 3, 5, 12].

Creative Media Production Social Machines
In [7] we describe *creative media production social machines* as “a broad class of systems where:

- “humans may use a purely digital, or combination of digital and analogue methods, and a degree of creative effort, to produce media content;
- “the content is published and publicly accessible on the web;
- “a global audience may consume, curate and commentate on this content in technologically-mediated environments.”

When we realise that users of a system implicitly adopt one or more online personas across multiple systems (websites, applications, communities), and intermix their activities and interactions between systems whether the technology explicitly allows it or not, we can cease to consider them as users. Rather, they are *components*, without which the systems wouldn't function; they are part of a larger socio-technical picture.

MOTIVATIONS
At present, metadata about content creator personas, activities and interactions, not to mention the content they produce, is typically hosted by a central content-host site (like YouTube). Content-host sites, whether they host the same types of media or not, typically do not allow cross-site interactions, or easy transfer of data from one site to another.

Reliance up on these silos is disempowering to content creators. Their success is usually hinged around their

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

metadata, for example how many videos they have uploaded, how many views their content has received, who their connections are. If a content-host site disappears, a large part of their creative portfolio and online networks may disappear with it. This is similarly the case if a creator's online profile is banned by a particular site (many creators on YouTube have problems with false copyright claims [11]).

They are also limited by the functionality provided by the site(s) they choose to use. Developers seeking to extend content-host site functionality are restricted by the APIs (or lack thereof) available. Despite the ability to share or embed content through social networks or personal websites, data about the content is still confined to the environment of a particular content host, and difficult to mash-up with content or data in other systems.

Availability of open web standards provides an opportunity to return ownership of data and metadata around content creation activities and interactions to the content creators themselves.

GOALS

Our goals therefore are to define a model for representing content creators and their content which will allow them to be referred to independently of the website on which content is hosted. We will implement this model using open web standards, namely linked data, to allow the creation of useful applications which can communicate with each other.

This outcome goes beyond simply facilitating the management of disparate content by its creator(s), but aims to allow them to have complete ownership of any and all data associated with a creative work, independently of their chosen creation and distribution platforms. They can then permit third-party applications to access the data needed to perform useful functions. Next, we describe some examples of when this will be useful.

Example scenarios

Changing content distribution platform: A successful amateur filmmaker publishes their work on YouTube. They have 100,000 subscribers, and millions of views. At some point, YouTube makes some visual or functional change to the website that the filmmaker doesn't like. Currently, the filmmaker's reputation and audience are tied to YouTube, so they have little choice but to tolerate the changes (even if the changes meant that they were more readily exploited by advertisers, or their content was less often recommended to people because of algorithm changes). Assuming instead that all content creators' data is available as linked data, and content host platforms which can consume linked data have been developed, they are free to publish elsewhere whilst retaining their reputation and links to their audience.

Creating a portfolio of creative work: A musician and sound designer who has recorded their own album (available on SoundCloud), collaborated with another musician to produce an EP (the other musician released it on BandCamp), made backing music for a short film (published by the film's director on YouTube) and created the soundscape for someone's animation (on Newgrounds) must manually collate this information onto a personal website to effectively build their portfolio. They have no web development knowledge and no resources to pay someone else to help. With all of the relevant data instead available as linked data, it is trivial for third-party portfolio-creation application to access and visualise it in a logical and attractive way. Metadata such as listens/watches, downloads or comments can be automatically integrated too.

Finding collaborative partners: In the current ecosystem, content creators must rely on personal networks or cold approaches to meet others to work with. For many who are just starting out, this is especially difficult. With a system that can access and interpret linked data to determine creators' skills, availability, interests and needs, partnerships can easily be automatically suggested, and perhaps initiated by the indication of automatically detected mutual friends.

APPROACH

We use ethnographical studies of content creators to inform the process of formally representing their identities and activities. The formal representation, an ontology for linked data, is in preliminary draft stages. We use this ontology to annotate content creator activities across a number of different platforms. We can then demonstrate the utility of this linked data by building applications that complement content creator activities regardless of the websites they use.

Studies

We interviewed 39 amateur content creators, selected at random during two days of *VidFest*, part of the 2013 London ComicCon. We learnt that two thirds work collaboratively with others but none used specialist tools to facilitate collaboration or improve their creative process.

We also observed current behaviour of a small but diverse sample of ten content creators 'in the wild', with the goal of reaching a better understanding of their online identities, connections, activities and interactions. For the ten creators, 93 online profiles were discovered. Through these profiles, we discovered that many creators represent different versions of themselves online by adopting multiple personas. One persona may have many online profiles across different websites. Personas sometimes even interact with each other as though they are completely separate people. More details of this study can be found in [7].

It is important to note that we have no desire to attempt to disambiguate the identities of content creators, to connect their online personas to an offline individual behind them, or to otherwise infringe upon web users' right to pseudonymity or anonymity. Rather, we argue that it is crucial to acknowledge when, how and why creators use different presentations of themselves, in order to improve the systems they use.

Ontologies

Many existing ontologies are suitable to some degree for representing aspects of digital content creators. We summarise the key ones here:

- FOAF: Friend-of-a-Friend can be used to represent connections between people [6]. We extend this to allow more fine-grained types of connection.
- SIOC: Semantically-Interlinked Online Communities provides core concepts for linking textual exchanges of messages to each other and to their creators [2].
- OntoMedia: this provides generic content description for various media types, and allows descriptions of relationships between different media types [9].
- PROV-O: The W3C Provenance Ontology allows annotation of an activity, the inputs and outputs, and the roles that people play [10]. We will extend this to provide more domain-specific detail, where the activity described is the 'creation process' of a piece of media content.

CONCLUSION

We have outlined some of the limitations experienced by content creators in the current user-generated content ecosystem. Our preliminary studies show that the complex and nuanced identity behaviours of content creators makes it imperative to consider their experience within creative media production social machines, rather than on individual websites and services, in order to design and develop improved systems.

REFERENCES

1. Berners-Lee, T. Weaving the Web: The Past, Present and Future of the World Wide Web by its Inventor. London. (1999). Texere.
2. Breslin, J., et al. SIOC: an approach to connect web-based communities. *International Journal of Web Based Communities*, 2(2), 133-142, (2006).
3. Byrne Evans, M., O'Hara, K., Tiropanis, T., and Webber, C. Crime Applications and Social Machines: Crowdsourcing Sensitive Data. In *WWW 2013 Companion*, Rio de Janeiro (2013). ACM.
4. Dalton, B. Pseudonymity in Social Machines. In *WWW 2013 Companion*, Rio de Janeiro, (2013). ACM.
5. De Roure, D., et al. Observing Social Machines Part 1: What to Observe? In *CHI'13*, Paris, 2013. ACM.
6. Graves, M., Constabaris, A., and Brickley, D. Foaf: Connecting people on the semantic web. *Cataloging & Classification quarterly*, 43(3-4), 191-202, (2007).
7. Guy, A. and Klein, E. Constructed Identity and Social Machines: A Case Study in Creative Media Production. In *WWW 2014 Companion*, Seoul (forthcoming). ACM.
8. Hendler, J. and Berners-Lee., T From the Semantic Web to social machines: A research challenge for AI on the World Wide Web. *Artificial Intelligence*, 174(2):156–161, (2009).
9. Jewell, M.O., Lawrence, K.F., Pruegel-Bennet, A. and schraefel, m.c.. Annotation of Multimedia Using OntoMedia. In *First International Workshop on Semantic Web Annotations for Multimedia (SWAMM)*. Edinburgh, 2006.
10. Lebo, T., et al. PROV-O: The PROV Ontology. Working Draft, W3C, 2012.
11. Power, M. "Looks like this is goodbye.." [Facebook update] Retrieved from https://www.facebook.com/powerm1985/photos/a.267403889986893.65529.177637432296873/656515964409015/?type=1&theater on 16 February 2014.
12. Strohmaier, M. A Few Thoughts on Engineering Social Machines. In *WWW 2013 Companion*, Rio de Janeiro, 2013. ACM.

Chi sparks 2014

Make Your Own HCI!

Annika Geurtsen
LIACS, Leiden University
Niels Bohrweg 1, Leiden
a.geurtsen@umail.leidenuniv.nl

Bobbie Smulders
LIACS, Leiden University
Niels Bohrweg 1, Leiden
b.smulders@umail.leidenuniv.nl

ABSTRACT
This paper describes and discusses an explorative research into the design and functionality of a game controller that can be built entirely from paper and without the use of electronic components.

Author Keywords
Analog controller; User satisfaction; AR interaction.

ACM Classification Keywords
H.5.2: Information Interfaces and Presentation (e.g., HCI): User Interfaces

INTRODUCTION
This project aims to provide a starting point for exploring the designing and building of hardware (in this case a game con- troller) without using electronic components. This paper dis- cusses a game controller free of electronic components and existing entirely out of paper, that can be placed on a plat- form in order for users to play games. The idea originated from a discussion on the seemingly high threshold building custom hardware (containing electronic components) poses on non-technical people interested (people that we were familiar with) in building and using custom hardware.

Even though platforms exist that encourage people to start working with hardware, like Arduino [1], Dwengo [3], and Phidgets [6], these platforms are still using electronic components, thus requiring the user to possess a certain amount of knowledge about electronics before being able to build their own custom hardware. The aim of the paper game controller is to enable non-technical users to quickly design and build a functioning piece of hardware (a controller) without the need of prior knowledge on electronics and thus lowering the threshold for non-technical people to start building their own hardware [5].

THE WORKINGS OF THE PAPER GAME CONTROLLER
The game controller described in this paper is made out of paper and free of electronic components, enabled by the use of AR markers (Figure 1). The controller consists of a paper box the size of a shoebox (but it could have any size) that can be mounted with 3 different types of buttons: a turning knob, a toggle switch and a slider (Figure 2 & Figure 3).

Note that these buttons are just options for buttons. In theory any type of button can be created as long as the software is able to recognize a distinct motion.

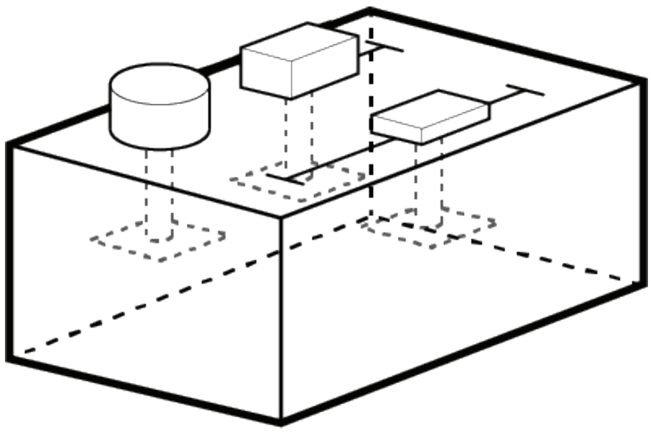

Figure 1. The paper game controller existing of a (shoe sized) box that is mounted with 3 types of buttons (turning knob, toggle switch and slider). Each button type can appear more often on the game controller or not at all depending on the wishes of the user.

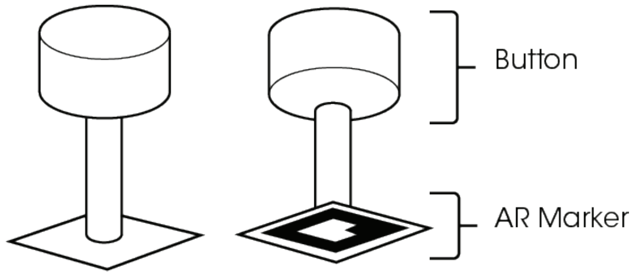

Figure 2. Anatomy of the button

Each button is provided with an AR marker that in turn is registered by a webcam connected to a computer that runs custom code in Processing (Figure 5). The button is recognized by the system through motion of the AR marker: the turning knob button has to turn, the toggle switch has to slide between a distance of 2 to 3 cm and the slider has to slide between a distance of 4 or larger.

The AR marker is placed facing downward so that a webcam can be placed below the controller. In doing so, the user can use and manipulate the buttons without obstructing the view of the AR markers for the webcam.

The webcam is installed inside a pedestal to create distance between the webcam and the AR markers. This is done in

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

order for the webcam to read the AR markers more precisely. Furthermore a light is installed next to the webcam to allow for more contrast between the black and white of the AR markers resulting in more precise vision of the webcam (Figure 5).

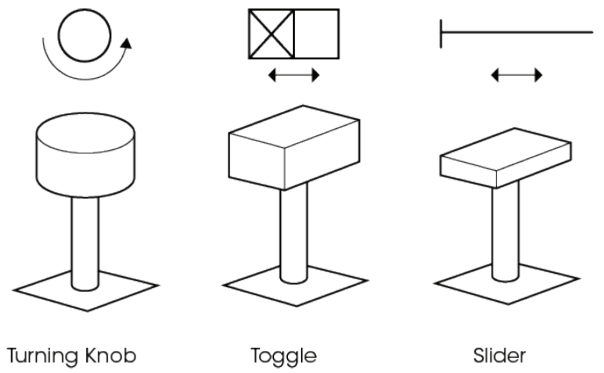

Figure 3. Types of buttons

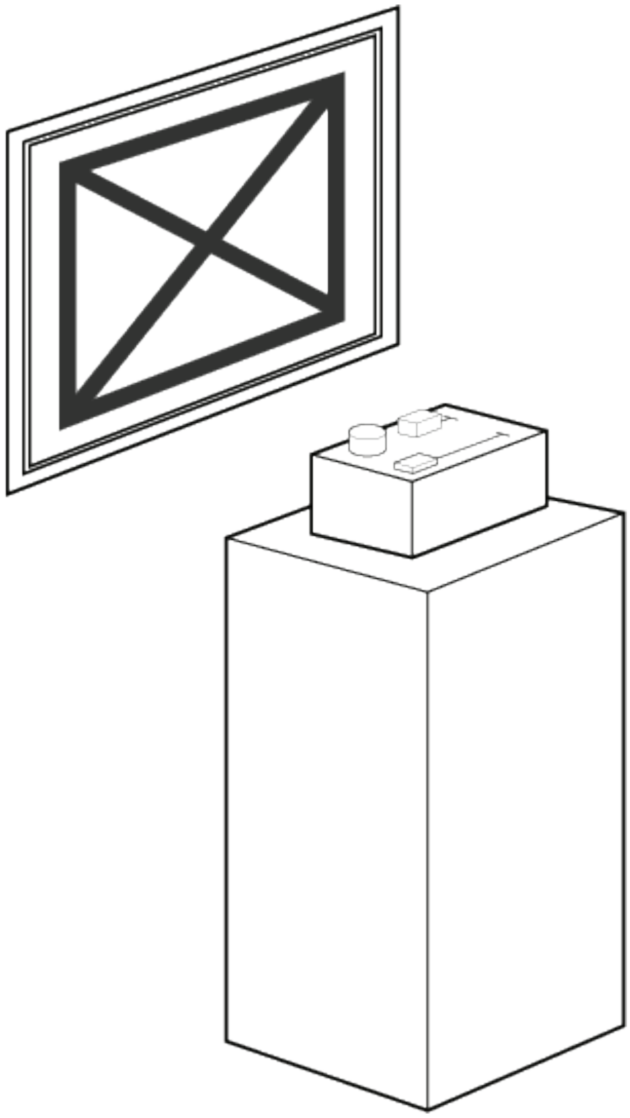

Figure 4. Set up of the game controller for gaming. The paper game con- troller is placed upon a pedestal and the user will use the game controller to control a game on the screen

The game that can be played in this case is a game created specifically for the paper game controller. The game uses a two-dimensional physics engine that can be controlled by the three types of buttons. The game screen is filled with balls.

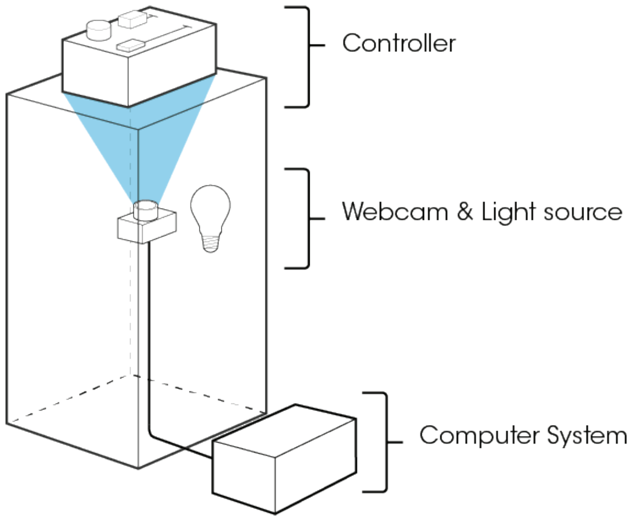

Figure 5. Set up of the paper game controller with the webcam (aided by a light source) that reads the AR markers, and a computer system that interprets the button types

A slider button can be used to control the overall gravity (be- tween non-existent and earth-like), a turning knob can create a local gravity field at the location of the knob and a toggle switch can switch all turning knobs from attracting the balls to repulsing the balls.

DISCUSSION — KEY CHALLENGES

The choice of material for the game controller is paper, due to its flexibility (it can be folded, cut and pasted) and the familiarity of users with the manipulation (e.g. cutting) of this medium. Unlike metal or plastics, paper is a medium that is easy to manipulate in a home environment without the need of special skills (e.g. metalworking). One of its disadvantages is that the possibility of the paper controller breaking down (e.g. due to rough handling) is larger than when the game controller would have been made out of wood or plastics. On the other hand building a controller using paper requires less time and skill in comparison to building a controller made out of wood. The scope of this project is to enable users to quickly and easily build their own game controller, not to create a lasting controller. Therefor paper as a medium to build the game controller suits this project better than wood, metal or plastics.

The main challenge presented by the paper game controller is the absence of electronic components in the game controller itself. This means that a standard, direct communication between the game controller and a computer system that controls the game is not possible and that an indirect means of communication between both systems must be set up.

Various platforms exist that enable communication between the analog game controller and the computer system on which the game is played. Kinect [4] states that nothing else but you should be the controller. A motion sensor registers the users movements and creates a digital skeleton based on distance measurements. Even though this is a very nice method for creating a controller without any electronic parts (stating that the user is the controller and that the sensor, though part of the communication to the computer system, is not), it does not allow the user to build his own custom hardware.

An example of natural feature tracking as a way of creating an analog controller is SketchSynth [7] in which a user can draw figures on paper and by doing so creates functioning buttons that can be used to perform tasks. Once drawn, the controller sends Open Sound Control messages to a synthesizer running in Pure Data. Even though both projects emphasize the ease of interaction (through drawing in SketchSynth and using paper to build the analog game controller), the paper game controller differs from SketchSynth in the sense that rather then having buttons drawn on a 2 dimensional surface, it is a 3 dimensional controller that needs to be built. In our opinion the paper game controller approximates the concept of building hardware more than the SketchSynth, because in the Sketch-Synth project a controller is being drawn and not really being built.

The reason for using AR markers is that they can be printed or (if necessary) even drawn on paper, without the drawing itself becoming the button. A webcam detects the AR markers and feeds the information to the program. This makes the AR marker completely free of electrical or digital components and easy to make. Also AR markers are simple in use, since they just need to be cut and pasted on the bottom of a button.

Another project venturing into simplifying the use of hard-ware is d.tools.[2] It is a hardware and software system that was built to support design thinking rather than implementation tinkering. The set up of electronic components (physical controllers, sensors and output devices) is simplified. In this sense d.tools shares an identical aim as the paper game controller. However it does make use of (simplified) electronic components, whereas the paper game controller tries to eliminate electronic components in the game controller all together. Next to that d.tools aims to create a platform for rapid prototyping of devices, so that designers can experiment and monitor physical user interfaces, whereas the paper game controller is aimed at a broader audience, focussing on the building of a game controller.

A project closer to the paper game controller is VoodooIO [8]. It is a physical interface consisting of a flexible sheet that allows for buttons to be aggregated and organized into it. VoodooIO is similar to the paper controller in that the intention of the project is to overcome the obstacles that prevent hardware interfaces from being easily appropriable. Also it uses a modular system to distribute buttons (or atomic units of control). However this project uses direct communication between the computer and the controller.

FUTURE WORK

Possible future work for the paper game controller include the design of different forms of boxes for the body of the paper game controller. This project focussed on using a box similar to a shoebox, but the body of the controller does not have to be limited to the shape of a shoebox. In fact a different shape of body might allow for different functionalities. The only restriction that is to be applied is that the webcam has clear vision on the AR markers.

A similar concept accounts for the design and functionality of buttons. In this paper three types of buttons are discussed: the turning knob, the toggle switch and the slider. This leaves room for investigation of other types of buttons, for instance pressing buttons.

Also the type of game could be investigated. Existing games could be implemented, so that users can play these games with their own custom build controller or games could be de- signed to specifically fit the paper game controller software system. Aside from games, other types of systems could be explored, such as rapid prototyping interfaces.

ACKNOWLEDGMENTS

We thank Amalia Kallergi and Fons Verbeek for providing helpful comments and support throughout the project.

REFERENCES

1. Arduino – homepage. http://www.arduino.cc/
2. HCI at stanford university – d.tools: Enabling rapid prototyping for physical interaction design, 2006. http://hci.stanford.edu/research/dtools/
3. Dwengo: gets you started with micro controllers. http://www.dwengo.org
4. Xbox – kinect overzicht. http://www.xbox.com/nl-NL/kinect .
5. Korukonda, A. Personality, individual characteristics, and predisposition to technophobia: some answers, questions, and points to ponder about. Information Sciences 170 (2005), 309 – 328.
6. Phidgets inc. - unique and easy to use usb interfaces. http://phidgets.com .
7. SketchSynth drawable user interface by Billy Keyes. http://www.creativeapplications.net/openframeworks/sketchsynth-drawable-user-interface-by-billy-keyes/
8. Lancaster university computing department – voodooIO: Pin & play interface architecture. http://hci.stanford.edu/research/dtools/

Chi sparks 2014

Digicoach: a digital coach for self-tracking athletes

Antoine Hogenboom
Researcher University of Applied Sciences Amsterdam & Info.nl
Sint Antoniesbreestraat 16, Amsterdam
antoine@info.nl

ABSTRACT

The Quantified Self community is growing and life logging is becoming mainstream. More and more people are quantifying themselves using tracking devices. In existing devices there is no coaching functionality, the collected data is analyzed and interpreted by the athlete himself. For runners, we developed a digital coach that is able to provide feedback in an intuitive way, without interrupting the athlete's running flow. This is done by giving vibration pulses in combination with visual led feedback. The digital coach is built into a wristband with GPS module, vibration motor and several LEDs. The next step is to give the digital coach a personality according to the DISC coaching model. This means that the digital coach can be either dominant and strict, or just relax and absent. From the data collected the digital coach should identify which style gives the best results for the athlete.

Author Keywords

Quantified Self; Coaching; Haptic feedback; Feedback loop; Self-learning systems; Running

ACM Classification Keywords

H.5.m [Information interfaces and presentation (HCI)]: Miscellaneous

INTRODUCTION

Self-tracking, lifelogging or quantified self is the phenomenon by which people measure their own behavior or activities using tracking devices [1]. This way they are trying to get insights into their own lifestyle in order to be able to improve it. The most popular category of trackers is fitness and sports related.

The iterative process of measuring, analyzing data and modifying behavior is called a feedback loop [2]. All these measurements delivers a huge amount of data, but data is not necessarily equal to insights. Many existing applications and fitness trackers only measure the performance of the athlete, but do not give any feedback. The user must interpret the data himself. Moreover, this data analysis is done only after an activity. Especially for athletes, it is desirable that they get immediate feedback on their performance during exercise.

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

GOAL

The goal of this research project is to add a digital coach to the feedback loop. A coach that analyzes the generated data, interprets it and provides feedback based on this. This feedback should be given not only after, but also during exercise, so the athlete can immediately alter his behavior.

LIVE FEEDBACK

We've developed a wearable coaching module for runners. using a combination of tactile sensors and a GPS module. A rich sensing module is created that can be used for coaching runners. The device is a wristband with GPS module, vibration motor, and LEDs.

The wristband maps the running behavior of athletes using GPS. The collected data is displayed in an online dashboard . The feedback to the athlete is given in an intuitive way while running, using LEDs and vibration motors. On first use, the digital coach will do a baseline measurement to determine the basic level of the athlete. Based on this, the coach sets goals, for example to run a certain distance or reach a certain average speed. These goals will be reassessed before each run. The digital coach can give three 'commands': go faster, go slower, or " you're doing fine, keep it up." These commands are given through vibration impulses.

FUTURE WORK

The frequency in which the digital coach provides feedback and the intensity in which it is given, is currently fixed. The goal of the project is to develop a personal coach that adapts the frequency and intensity of vibration to the athlete. The next step in the research is to give personality to the digital coach. One athlete benefits from a strict coach, the other just needs one laid back. The characteristics of the digital coach will be developed based on the DiSC model, or the Five Factor Model. These models distinguish different types of coaches. This means that the digital coach can be either dominant and strict, or just relaxt and absent. From the data collected on the athlete, the coach should identify which style fits best to the athlete, in order to really develop a personal coach.

Future research will focus on investigating how a digital coach can adapt to the athlete, so that over time a personal coach will be developed. We will investigate how we can achieve matchmaking between the coach and the athlete.

ACKNOWLEDGMENTS

This research was partly funded by the CREATE-IT knowledge center for applied research of the University of Applied Sciences Amsterdam and Internet agency Info.nl Amsterdam

REFERENCES

1. Rivera-Pelayo, Veronica, Zacharias, Valentin, Muller, Lars, & Braun, Simone. (2012). "Applying quantified self approaches to support reflective learning." *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, Vancouver, British Columbia, Canada.
2. Goetz, Thomas. (2011). *Harnessing the Power of Feedback Loops*. Retrieved 21/02/2012, from http://www.wired.com/magazine/2011/06/ff_feedbackloop/all/

chi sparks 2014

Iris Intranet: Enhancing the Way People Work

Arthur Turksma
triptic BV
PO Box 4279, 5604 EG Eindhoven
a.turksma@triptic.nl
+31 40 2463040

Jos Rouw
triptic BV
PO Box 4279, 5604 EG Eindhoven
j.rouw@triptic.nl
+31 40 2463040

ABSTRACT

In this paper, triptic BV describes Iris Intranet, the intranet platform that enables people to deliver better business performance thanks to its HCI design. Iris Intranet is one of the winners of the worldwide Intranet Design Annual 2014 organised by the Nielsen Norman Group, receiving praise for the appealing design: "One of several traits that differentiates Iris, the triptic intranet, from other intranets is that it actually looks like fun to use." [4]

Author Keywords

demo; usability; collaboration; intranet; system integration

ACM Classification Keywords

H.5.3. Information interfaces and presentation (e.g., HCI): Collaborative computing.

INTRODUCTION

An intranet that combines up-to-date, business-critical information with flexible collaborative tools, thus enhancing the way people work. That is Iris Intranet. The approach behind her award-winning design and process-oriented system integrations are explained in this paper. Also, plans on how to make Iris more context-driven are discussed.

THE MAIN OBJECTIVE: ENHANCING PERFORMANCE

The main objective of Iris Intranet is to enable people to work smarter, faster, better. By providing employees with all the essential business information they need, in one place. No more expensive minutes of searching unfriendly systems: Iris serves information in a dashboard that employees can tailor to their personal wishes. Other platforms currently available are mostly social intranets, aiming to stimulate knowledge sharing, but Iris does more. Combining process-driven information with flexible collaboration tools, Iris helps everyone in an organisation gain more insights in less time together, wherever they are. That's why Iris calls herself *the intranet that actually works*. Having seen the light in 2013, Iris has arrived at a good time. The McKinsey Global Institute, for instance, found that this kind of technology "could raise the productivity of interaction workers by 20 to 25 percent." [1].

THE DESIGN APPROACH

Working with usability methods as proposed by Krug [3] and (service) design thinking methods as discussed by Schneider and Stickdorn [5], and benefiting from years of web design and system development experience, triptic developed Iris Intranet into a platform geared to people and in the service of business processes.

The design of Iris Intranet is highly user-centered. Mapping out employee journeys following Schneider and Stickdorn [5], the (most probable) needs of people in a general workforce were identified and used as a starting point for the design, which was repeatedly subjected to usability testing to make sure people can work with Iris intuitively (i.e. 'without thinking') following the "Don't make me think" user perspective coined by Krug [3]. One key outcome of this designing for intuitive use is the so-called Swiss Army Knife, a toolbox that is always present wherever users go on the intranet and allows them to fulfil key tasks quickly. To make sure that employees can work with Iris on any device, she was designed responsively.

Figure 1. Screenshot Iris Intranet.

The result of the design approach was rewarded by research consultancy Nielsen Norman Group from Silicon Valley. In their worldwide Intranet Design Annual, they named the Iris that triptic works with as one of the 10 best intranets of 2014. "One of several traits that differentiates Iris, the triptic intranet, from other intranets is that it actually looks like fun to use." say the Nielsen Norman Group's consultants [4].

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

ENTERPRISE SYSTEMS INTEGRATED

To facilitate employees' information needs within the processes of their business context, Iris integrates with enterprise systems. Examples may be project management systems and customer relationship management systems. From there, Iris serves up-to-date business-critical information in bitesize chunks in widgets on the dashboard. This way, the dashboard shows employees the current state of affairs at a glance. Such process-related information will make the intranet more attractive to managers, whose presence is very likely to stimulate activity amongst employees [2].

MEETING REQUIREMENTS FOR SUCCESS

Haan et al. [2] from the Dutch consultancy Evolve studied the use of internal social media in 76 Dutch organisations with more than 50 employees. They formulated four requirements for success (which has to be defined in terms of for instance activity, participation, Return On Investment), namely use, goal orientation, formal organisation, and embedding in processes. With the seductive design of Iris Intranet – it "actually looks like fun to use" according to Nielsen Norman Group [4] – and goal oriented integration of process-based information, Iris enables organisations to quickly cover three of these four requirements. Thusly, she sets an example in the field of intranets and enterprise social media, increasing the relevance for and impact on business performance. The remaining requirement, the formal organisation of the internal community on the intranet, is a change management initiative, most likely to fall in the hands of HR, communications and senior management.

THE NEXT AIM: TO BE MORE CONTEXT DRIVEN

Currently, the architects behind Iris Intranet are studying ways to make the intranet even more context-driven, mostly based on the employees' location. Use will be made of technology which is a driving force of the Age of Context, a term coined by Scoble and Israel [6] to describe how current technological developments allow for extensive personalization in a wide range of products and services.

CONCLUSION

The concept of what an intranet is is changing. Iris is a pioneer in this particular shift, because of its HCI design. The result is one platform with both collaborative social tools and, thanks to system integrations, business-critical information. A combination that enables people to work smarter, faster, better. Anytime, anywhere, thanks to the responsive design. To facilitate the needs of employees in the workforce, Iris is based on a highly user-centered design approach containing certified usability methods and service design thinking methods. People can work with Iris intuitively, and are seduced to do so by the appealing design for which the intranet received an international award. In the future, Iris will become more context-aware, further enabling people and organizations to perform better.

REFERENCES

1. Chui, M., Manyika, J., Bughin, J., Dobbs, R., Roxburgh, C., Sarrazin, H., Sands, G. & Westergren, M. (2012). *The Social Economy: Unlocking Value and Productivity through Social Technologies*. Retrieved from: http://www.mckinsey.com/insights/high_tech_telecoms_internet/the_social_economy on March 15th, 2014.
2. Haan, P., Lasseur, A., Dechesne, Y. & Hazelaar, K. (2014). *Interne Social Media in Nederland: de Stand van Zaken*. Retrieved from: http://www.evolve.eu/index.php/Wonder-interne-social-media-in-nederland-veel-enthousiasme-ruimte-voor-professionalisering/ on March 15th, 2014.
3. Krug, S. (2000). *Don't Make Me Think! A Common-Sense Approach to Web Usability*. Berkeley, CA: New Riders.
4. Pernice, K., Schade, A., Caya, P. and Nielsen, J. (2014). *Intranet Design Annual 2014: The Year's 10 Best Intranets*. Fremont, CA: Nielsen Norman Group.
5. Schneider, J. & Stickdorn, M. (2011). *This is Service Design Thinking*. Hoboken, NJ: John Wiley & Sons Inc.
6. Scoble, R. & Israel, S. (2013). *Age of Context: Mobile, Sensors, Data and the Future of Privacy*.

chi sparks 2014

ixi-Play, a robot buddy for young children

Bart Dirkx
WittyWorX LLP
Brouwer 4, Eersel
bart.dirkx@wittyworx.com

Ruud van de Aalst
WittyWorX LLP
Brouwer 4, Eersel
ruud.van.der.aalst@wittyworx.com

ABSTRACT

Current tablet computers are so easy to use, that even children can operate them. But is a virtual environment operated by a touch screen really the optimal form for a child to develop new skills? In this paper, WittyWorX describes ixi-Play, a robot buddy for young children to play physical games with in the real world. This buddy can see, listen, feel, show emotions and move life like enabling young children to explore and play games in the physical world we live in. Communication with the ixi-Play platform takes place by speech recognition, image recognition and touch. ixi-Play not only offers rich interaction, the character also shares time and space with the child offering a higher level of engagement than screen based media.

Author Keywords

Education; Robotics; Personal Robotics; Social Robotics; Children; Gaming; Human Machine Interaction

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION

The ongoing development in semiconductor technology has enabled cheap microprocessors and sensors propelling the market of computer technology. Nowadays desktop computers and laptops are accompanied and in many cases replaced by smartphones and tablet computers. In combination with powerful operating systems e.g. Android™ or iOS™ that can be controlled with only your fingertips, these devices have become the new way of how we interact with computer technology. In fact these devices are so easy to use, that even children can do it. This incentivized many developers to develop children's Apps opening a whole new market for tablet- and smartphone-based games. So using tablet computers and smartphones has literally become child's play.

Parents love to see their toddlers master these devices, believing it will boost their child's IQ so they will become geniuses. Research however has shown that the opposite is true. Frequent screen use can have several negative side effects e.g. sleep disorder [1], attention disorder and delayed language learning, whereas the potential positive effects are low to non-existing. That is why the American Academy of Pediatrics advised to avoid screen-based media below the age of two [2].

Research has shown that we learn better and faster if we combine cognitive with physical interaction [3]. Tablet computers and smartphones offer only audiovisual feedback and interaction via a touch screen by tapping and swiping. The images on screen may represent a real character, but lack the physical detail of their real life version. They remain a 'picture behind glass' and that is one of the reasons why people still buy action figures and plush toys of their favorite characters. Even though they cannot move or interact at all, children are often more attached to them than to the virtual character.

ROBOT TECHNOLOGY

What if we could bring characters to life with robot technology? Although robots have been around for decades, they have not seen the tremendous development that computers have. Of course with the development of sensors e.g. camera modules, microphones or accelerometers, and advanced mobile operating systems e.g. Android™ and iOS™ supporting speech recognition and image recognition Human Machine Interaction has improved [7]. The missing element for natural interaction is natural motion. Most robots move 'robotic' i.e. simple linear motion, one joint at a time while making a buzzing noise. This limits convincing human interaction and that is why WittyWorX has developed a motion module that enables fast, smooth, silent and agile motion in six degrees of freedom [4].

Another reason why robot technology has not seen the same development as computers is because of their price. Advanced robots like Asimo™ and NAO™ can do complex operations, but are far too expensive for large-scale adoption by consumers. The reason for this is that they use high-end modules and components e.g. motors and gearboxes. Toy robots are affordable, but because they use low-end modules they are mostly noisy or slow and have limited functionality. Therefore they do not offer the same user experience as advanced robots. Mechanical components do not scale the same way as electronic components, because they do not become faster and cheaper by making them smaller. The way to develop an affordable robot is by further integration, designing out the noisy or expensive components and leaving out the unnecessary elements.

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

IXI-PLAY PLATFORM

This motion module forms the basis for ixi-Play, an intelligent platform that can bring characters to life (figure 1). Ixi-Play holds a camera for image recognition, two microphones for hearing, a multitouch sensor for feeling touch, two small color displays for showing eye animations, a flexible body for lifelike motion, two speakers for music, speech and sound effects. The platform also offers WiFi and Bluetooth for wireless communication with other devices and the Internet.

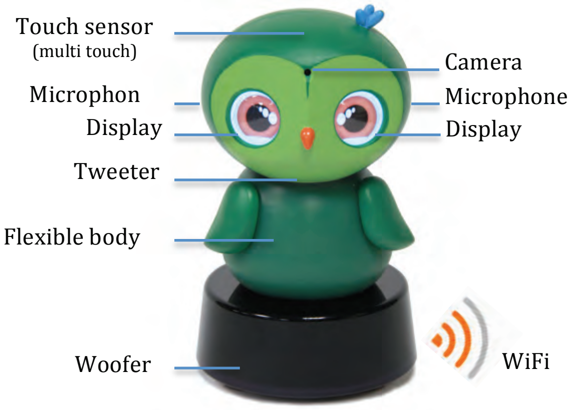

Figure 1. ixi-Play platform with sensors

The ixi-Play platform holds a powerful computer capable of doing on the fly image recognition e.g. face detection, color or card recognition. This enables ixi-Play to recognize objects in the real world. The platform runs Android™ operating system making it easy to program. It also allows other devices e.g. tablet computers to interface even if they run a different operating system e.g. iOS™ or Windows™.

EMOTIONS AND EXPRESSIVENESS

Ixi-Play's motion module together with the flexible body, eye animations and sound effects enables the character to have a high level of expressiveness and show convincing emotions (figure 3). As for animated movie characters, also for physical characters, fluent and silent motion appears to be crucial as it amplifies user experience. Masahiro Mori researched this aspect as published in his Uncanny Valley paper [5]. Figure 2 shows the added value of motion in the perceived affinity of the character versus the user.

In this figure ixi-Play positions itself just before the uncanny valley as it has a higher affinity than a humanoid robot, but is far from resembling a zombie.

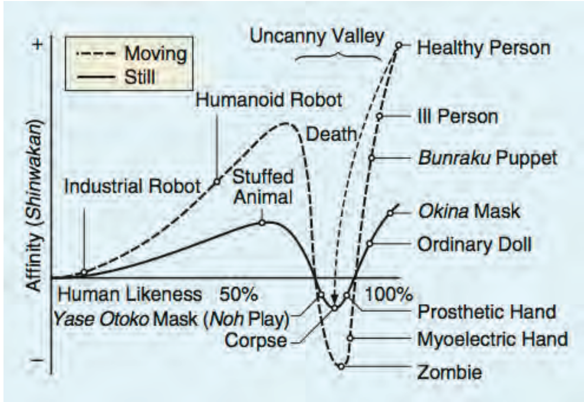

Figure 2. The presence of motion steepens the slopes of the uncanny valley [5]

APPLICATIONS AND ACCESSORIES

An important aspect of Ixi-Play is that it enables children to play games in the real world with physical objects and interaction. This can be both guided and open-ended play. Examples of possible games are: parroting, colors & shapes, animal sounds, storytelling, connect four, music & dancing and dressing up (figure 4). These games can be played together with accessories like flashcards or colored blocks that can be recognized by ixi-Play as well as clothes.

While playing the child learns how to manipulate objects and interact with them. In the mean time ixi-Play acts like a peer, so the child plays the games against or with ixi-Play. Ixi-Play interacts by watching, listening, speaking, moving and making sounds.

Because ixi-Play is not capable of moving objects, the children have to do all the physical operations, which is good because this improves not only their physical skills, but also makes them learn faster. Games can be educational or just fun. The educational content is mainly determined by the game itself. Ixi-Play makes it fun to learn without the need of an adult.

Besides being a game companion, ixi-Play also serves as a baby-monitoring device where video is streamed to a tablet or smartphone via WiFi or as timekeeper. Because ixi-Play runs Android he can speak and recognize over 40 languages, which makes him a great aid for learning new languages. Research by Kanda et. al. [6] has shown that the learning effect improves when the child develops a relationship with the robot.

Figure 3: Examples of emotions of the ixi-Play platform

Figure 4: Examples of games with ixi-Play

UNIVERSITY RESEARCH

Leiden University has used ixi-Play for their research on how children learn [8]. The main goal of the study was to examine whether the robot could be a useful tool in dynamic testing and whether it was able to improve children's performances on a series completion task (Towers of Hanoi) and a complex reasoning task. Their research indicates that the robot can be of additional value for a dynamic assessment. The children were able to learn from the graduated prompts training which was given by the robot. Moreover, they were able to change their strategy to more analytic behaviour. Once the robot is able to register all actions by itself, testing with the robot will be a promising method to have a standardized dynamic test, which will not be a large time consumer for the test leader.

Eindhoven University of Technology has carried out research on the added value of motion to a human interface robot [9]. They measured motivation and attention by looking at verbal and non-verbal communication of the children. Their research showed that movement adds to the user experience, but may also be distracting if the robot is not part of the game. They suggest involving the robot in the game to make use of the attention it draws from the child.

Delft University of Technology has researched if children are able to recognize emotions from a robot by just using eyes, posture and sound [10]. The results were that this is true for all children tested (4-6 years of age). Tests with parents show that even children below 2 years of age can recognize these emotions. Delft's research also showed that interaction with an avatar (virtual representation of the character) is much different from interaction with the real, animated version of the character. Typical behavior with the avatar was that children tried to draw attention by tapping on the screen and shouting, whereas with the real character they had immediate contact by touching the head and talking.

CONCLUSION

With ixi-Play WittyWorX has developed an interactive platform that is affordable and easy to use offering a level of engagement that outperforms tablets or smartphones by enabling children to play games in the real world with tangible objects. University research has shown that this platform contains the necessary elements to do effective and efficient research on both Human Machine Interaction and child learning.

Future work contains adding games and accessories to the platform as well as improved image recognition and intelligence. Furthermore the platform needs to be tested in a broader setting to verify the long-term effects and benefits on children.

ACKNOWLEDGMENTS

We thank all children, parents, University researchers and volunteers, who helped us in creating and testing ixi-Play. Their input and advice has helped us in making a platform that fits their needs.

REFERENCES

1. Michelle M. Garrison and Dimitri A. Christakis, "The Impact of a Healthy Media Use Intervention on Sleep in Preschool Children." *Pediatrics*, originally published online (2012), DOI: 10.1542/peds.2011-3153.
2. AAP Council on Communications and Media. "Media Use by Children Younger Than 2 Years, Policy Statement." *Pediatrics*, DOI:10.1542/peds.2011-1753
3. Hoenig, K, Sim, E.-J., Bochev, V., Herrnberger, B. Kiefer, M. (2008). "Conceptual flexibility in the human brain: Dynamic recruitment of semantic maps from visual, motion and motor-related areas." *Journal of Cognitive Neuroscience 20*: 1799-1814
4. WittyWorX ixi-Play. http://www.ixiplay.com
5. Mori M., MacDorman K.F., Kageki N (2012). "The Uncanny Valley." *From The Field.*
6. Kanda T., Hirano T., Eaton D, Ishiguro H. (2004). "Interactive Robots as Social Partners and Peer Tutors for Children: A Field Trial." *Human-Computer Interaction, 2004,* Volume 19, pp. 61–84
7. Thrun, S. (2004). "Toward a Framework for Human-Robot Interaction." *Human-Computer Interaction, 19*
8. Karlijn A. Nigg (2012). *Dynamic testing using a table sized robot. Assessing and improving a child's cognitive performances?* (Research Master Thesis Developmental Psychology). Developmental Psychology, Institute of Psychology, Leiden University
9. A. Mileounis, M. Boerhof, T. Martens, R. Jansen, E. de Vries, H. van Weperen (2013). *Robot with a Smile!* (Research Thesis Human Technology Interaction.) Industrial Engineering an Innovation Sciences, Eindhoven University of Technology
10. A. Düzenli (2012). *MyRo, Children's Robotic Companion*. (Research Master Thesis Design for Interaction.) Industrial Design, Delft University of Technology.
11. Asimo, Honda: http://asimo.honda.com
12. NAO, Aldebaran Robotics: http://www.aldebaran-robotics.com

chi sparks 2014

Interactive Visualization of Video Data for Fish Population Monitoring

Emma Beauxis-Aussalet
CWI
Science Park 123, Amsterdam
emmanuelle.beauxis-aussalet@cwi.nl

Lynda Hardman
CWI
Science Park 123, Amsterdam
lynda.hardman@cwi.nl

ABSTRACT

The recent use of computer vision techniques for monitoring ecosystems has opened new perspectives for marine ecology research. These techniques can extract information about fish populations from in-situ cameras, without requiring ecologists to watch the videos. However, they inherently introduce uncertainty since automatic information extraction is imperfect. To be accepted for scientific use, video analysis tools must support the analysis of the extracted information and of their uncertainty. Another challenge concerns the diversity of scientific interests. Ecologists have diverse research goals and information needs, for instance specific species, time periods, or locations. We present a visualization interface addressing these two challenges: providing information about fish populations as well as computer vision uncertainty; and enabling the exploration of specific subsets of the video data depending on user needs.

Author Keywords

Data Visualization; Uncertainty; Interaction Design; User Trust; Computer Vision Application.

ACM Classification Keywords

H.5.2. [User Interface]: Graphical user interfaces;

INTRODUCTION

The Fish4Knowledge project (fish4knowledge.eu) has continuously recorded video footage of coral reef fish from 9 underwater cameras during 3 years. This collection motivated the use of computer vision for automatically recognizing fish from different species, and monitoring the population dynamics. The original video collection is processed in 3 steps: sequencing of continuous video streams into 10-minute clips (for storage purposes), identification of fish amongst other objects, and recognition of fish species.

From an ecologist's perspective, each information processing step potentially introduces errors: from data collection (recording videos) and processing (recognizing fish), to interpretation (deriving facts from automatic fish counts). We developed visualizations and interaction designs for exploring these uncertainties. Our interface[1] discloses the data collection and data processing methods and their possible biases. It lets users walk through explanations of the video analysis processes and their uncertainty. The interface offers interactive visualizations of fish populations, e.g., to explore numbers of fish from specific species, locations or time periods. The uncertainties of the observed fish populations are also visualizable. Finally, users can annotate and share their findings.

ELICITING AND ADDRESSING USER NEEDS

Ecologists are not experts in computer vision systems and their technical concepts. The information about provenance and uncertainty needs to be comprehensive and sufficiently detailed, while remaining understandable. We investigated user information needs by interviewing ecologists and computer vision experts [1-3]. Regularly collecting user feedback along design stages revealed user needs throughout all stages, depending on users' expectations and knowledge of computer vision. User-Centered Design had limitations for introducing this novel technology: non-experts did not foresee all uncertainties inherent to the technology, nor appropriate metrics for interpreting them (e.g., ground-truth evaluation). Requirements from non-experts only would produce *incoherent design* as [5] notifies. Interviews of computer vision experts complemented the information requirements. Computer vision experts were particularly needed for supplying uncertainty metrics, and specifying *Uncertainty* and *Error Modeling* [4]: leveraging user support depended on experts' uncertainty evaluation. We identified 5 main information needs and 10 underlying uncertainty factors (Tables 1-2). Addressing such multi-factorial uncertainty issue is a key challenge in visualization research [6-8]: *"The R&D community must […] develop methods and principles for representing data quality, reliability, and certainty measures throughout the data transformation and analysis process"*[9]. Such methods and principles are not mature in our application domain: except [10] similar applications evaluated uncertainty with ecology-specific methods rather than computer vision methods (e.g., ground-truth evaluation). Our interface bridges this gap by developing visualizations of computer vision uncertainty accessible to non-experts. We introduce a novel visualization of ground-

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

[1] The prototype is online: f4k.project.cwi.nl

truth evaluation for non-experts (Figure 3), and an interaction design for exploring multiple uncertainty factors.

Information Need	Source	UI Tab
Watch Videos	Users	*Video* tab
Fish and species counts	Users	*Visualization* tab
Trends and Correlations	Users	*Visualization*, *Report* tabs
Uncertainty of Computer Vision system	Users, Experts	*Video*, *Video Analysis*, *Visualization* tabs
Uncertainty of Computer Vision results	Users, Experts	*Visualization* tab

Table 1. The main information needs.

EXPLORING VIDEO PROCESSING STEPS

The user interface tabs deliver manageable units of information and reflect the information processing sequence: data collection (*Video* tab), data processing (*Video Analysis*, *Extracted Data* tabs) and data interpretation (*Visualization*, *Report* tabs).

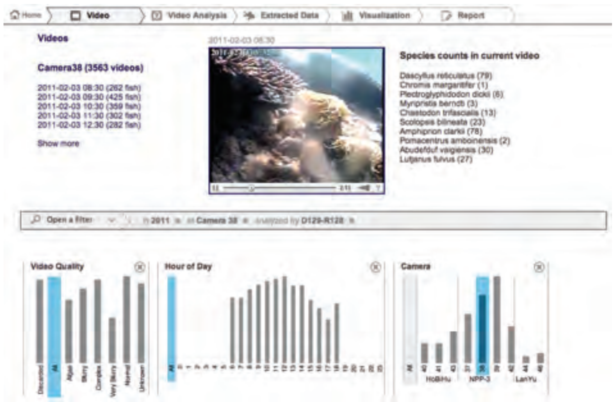

Figure 2. The *Video* tab.

The *Video* tab supports video browsing (Figure 2). It contains filtering functionalities for specifying the videos of interest (e.g., at specific location or time periods). Users can control the data collection conditions: which ecosystems are observed, with which field of view and image quality (e.g., lens biofouling, water turbidity).

The *Video Analysis* tab provides explanations of the video processing steps and visualizations of their uncertainty. It exposes the technical concepts needed for understanding computer vision uncertainty. The *Overview* sub-tab provides explanations of the main video processing steps. The *Fish Detection,* and *Species Recognition* sub-tabs provide visualizations of ground-truth evaluations (Figure 3). The *Workflow* sub-tab provides on-demand video processing. Users can request the analysis of specific videos (from user-defined time periods and cameras) with specific component versions (e.g., with the best accuracy for the species of interest). It serves either for processing videos that were not yet analyzed, or for experimenting with different versions of the video analysis components (e.g., to check robustness of observations).

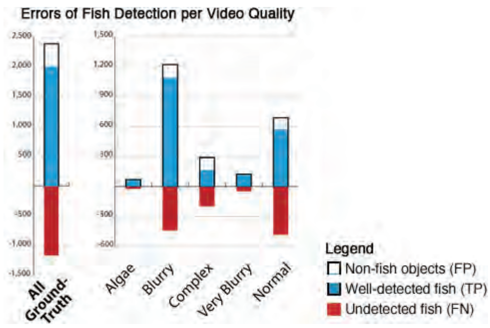

Figure 3. Visualization of ground-truth evaluation.

The *Extracted Data* tab provides an overview of the available video data and their properties (i.e., their dimensions). It shows all the characteristics of fish extracted from the video footage. It also explains the 4 main metrics provided for describing fish populations and their uncertainty: Number of Fish, Number of Video Samples (to check for missing videos), Number of Fish per Video Sample (to compensate for missing videos), and Number of Species. It helps understanding how fish populations can be monitored, and identifying the information relevant for particular studies.

VISUALIZING COMPUTER VISION UNCERTAINTY

The computer vision community uses well-accepted methods for evaluating uncertainty. They basically rely on a ground-truth: a set of images with manually identified objects, to which automatically identified objects are compared. A variety of metrics and visualizations can be derived from ground-truth evaluation. They are easily understood by experts but not by non-experts [1]. We adapted their visualization for non-experts. We reduced complexity by avoiding the use of advanced metrics (e.g., rates such as Precision/Recall), thresholds (e.g., ROC curves) and True Negatives (i.e., errors automatically discarded and having no impact on fish counts). [1] highlights that providing more details is likely to overwhelm users. The visualization exposes the proportions of items in the ground-truth, which is often omitted in more traditional evaluations. It indicates potential biases: the fewer the items, the higher the chance of error. The *Video Analysis* tab provides uncertainty visualizations (Figure 4) for each video analysis steps (fish detection and species recognition). Uncertainty is detailed for each video quality and each species, since different levels of errors indicate potential biases in the video data.

VISUALIZING COMPUTER VISION RESULTS

The *Visualization* tab (Figure 4) provides means to explore the video data, and the uncertainty due to missing videos, video quality or fish appearance quality. Videos can be missing due to camera maintenance, encoding errors, or unfinished processing queues. The quality of each fish appearance is measured using a *certainty score*. It indicates how much fish look like the fish model for its species. The higher the score, the more certain is the species recognition.

In Figure 4, Zone A contains the main graph. Zone B supports the adaptation of the main graph to specific user needs. Users can specify what the axes of the main graph represent. For instance, while the y-axis represents numbers of fish, the x-axis can represent their distribution over weeks of the year or hours of the day. Users can also select other types of graph (stacked chart or box plot). They provide additional information about the visualized fish population: e.g., the proportion of each species, or the variance of fish abundance. The selection of stacked charts or boxplot leads to the display of dedicated menus for adapting further the visualization. For instance fish counts can be stacked by species or by camera. Zone C contains filter widgets for both selecting datasets of interest, and overviewing datasets over several dimensions. Filter widgets are displayed on-demand. There are widgets for each dimension of the data, namely: Year, Week of Year and Hour of Day of fish occurrence, Camera, Species, Certainty Score, Video Quality and Software Version. A summary of the filters applied is provided in Zone B. To limit information overload, the default filters (e.g., all species, all cameras) are not mentioned in the summary. The widgets' histograms display the same metric as the main graph, and applied to the same dataset. E.g., in Figure 6 both the graph of Zone A and the histograms of Zone C display numbers of fish per video sample. Both use a dataset of fish detected by software D50-R52, occurring in 2011 at Camera 38, and belonging to all species, certainty score, video quality, week of year and hour of day. The *Camera* widget uses a dataset from all cameras, and highlights in blue which camera is selected.

The *Report* tab supports manual grouping and annotation of graphs created in the *Visualization* tab. Graphs can be added to and removed from a report. They can be annotated with a title and a comment. Users can download and upload reports in the form of text file containing a list of parameters. Reports can be saved and shared any file.

Our interaction design let users specify which data dimensions are relevant for their goal. Information of interest is displayed on-demand (open widgets in Zone C, change graph axes, display details in stacked charts and boxplots). Irrelevant information is not displayed (close widgets, switch back to simple graph). It provides both overviewing (widgets in Zone C) and detailed views (main graph in Zone A). It supports a wide range of data analysis goals, while limiting display cluttering and information overload. It addresses our design context with ecologists pursuing a variety of research goals, while being unfamiliar with video data.

CONCLUSION

Our interface allows ecologists to monitor fish populations using novel computer vision techniques. It allows the exploration of multiple uncertainty factors and aspects of the data. Early user feedback shows we achieved intuitive interactivity and easy to understand visualizations, although the dataset is unfamiliar to users. It allows preliminary data exploration for a variety of user goals, and the identification of uncertainties that may impact further information processing. This design informs and inspires other use cases dealing with open-ended data exploration, familiarization with novel data, or visualization of machine learning uncertainty.

REFERENCES

1. Beauxis-Aussalet, E., Arslanova, E., Hardman, L., & Van Ossenbruggen, J. A case study of trust issues in scientific video collections. *2nd international workshop on Multimedia Analysis for Ecological Data (pp. 41-46)*. ACM, 2013.
2. Beauxis-Aussalet, E., Palazzo, S., Nadarajan, G., Arslanova, E., Spampinato, C., & Hardman, L. A video processing and data retrieval framework for fish population monitoring. *2nd international workshop on Multimedia Analysis for Ecological Data (pp. 15-20)*. ACM, 2013.
3. Fish4Knowledge Project Deliverables: D2.1 User Information Needs; D2.3 Component-based Prototype and Evaluation; D6.6 Public Query interface. http://groups.inf.ed.ac.uk/f4k/deliverables.htm
4. Correa, C. D., Chan, Y. H., & Ma, K. L. A framework for uncertainty-aware visual analytics. In *IEEE Symposium on Visual Analytics Science and Technology (pp 51-58)*, VAST, 2009.
5. Norman, D. A. Human-centered design considered harmful. Interactions, 12(4), 14-19, 2005.
6. Johnson, C. R. and Sanderson. A. R. A next step: Visualizing errors and uncertainty. *Computer Graphics and Applications*, IEEE, 23(5): 6–10, 2003.
7. Chen, C. Top 10 unsolved information visualization problems. *Computer Graphics and Applications*, IEEE, 25(4): 12–16, 2005.
8. Johnson, C. Top scientific visualization research problems. *Computer graphics and applications*, IEEE, 24(4): 13–17, 2004.
9. Thomas J. J., & Cook. K. A. A visual analytics agenda. *Computer Graphics and Applications*, IEEE, 26(1): 10–13, 2006.
10. Hetrick, N. J., Simms K. M., Plumb M. P., Larson J. P. Feasibility of using video technology to estimate salmon escapement in the Ongivinuk River. Technical Report, *US Fish and Wildlife Service*, 2004

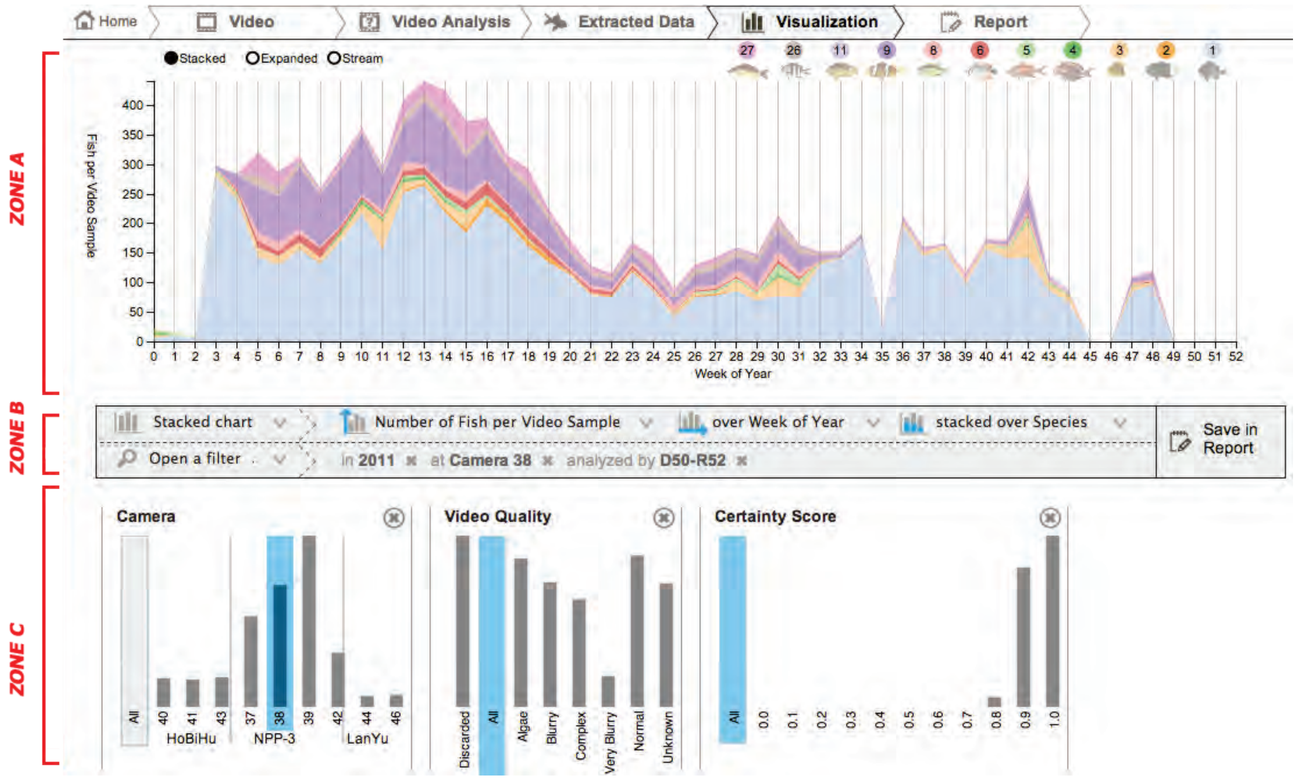

Figure 4. The *Visualization* tab.

Uncertainty Factor	Description	Source	Design Stage	UI Features
Data Collection				
Sampling Coverage	Camera recordings may not include all important areas and time periods.	Users	First Stage	Map of cameras (*Video Analysis* tab), Numbers of videos and fish per video (*Visualization* tab)
Image Quality	Recording conditions may impair the collected information (e.g., turbidity, fouling, video encoding).	Users, Experts	First Stage	Video browser, Classification of video quality, Filtering options (*Video, Visualization* tabs)
Field of View	Camera's fields of view may observe heterogeneous, incomparable ecosystems.	Users	Latest Stage	Video browser (*Video* tab)
Duplicated Individuals	Fish swimming back and forth are repeatedly recorded. Rates of duplication vary amongst *Fields of Views* and species swimming behaviors, thus producing biases.	Users	Latest Stage	*Evaluation method needs further research*
Data Processing				
Ground-Truth Quality	Ground-Truth items may be scarce, represent the wrong species or odd fish appearances.	Users, Experts	Mid-Term	Number of ground-truth items per species (*Video Analysis*)
Fragmentary Processing	Some videos may be yet unprocessed, missing or unusable.	Experts		Numbers of videos and fish per video (*Visualization* tab), On-demand video processing (*Workflow* sub-tab)
Fish Detection Errors	Fish may be undetected, and non-fish objects may be detected as fish	Experts		Ground-Truth evaluation (*Video Analysis* tab)
Species Recog-nition Errors	Species may be unrecognized, or confused with another.	Users, Experts	First Stage	Ground-Truth evaluation (*Video Analysis* tab)
Emerging Biases	Errors may be random (noise) or systematic (bias). Biases may emerge from combining data collection (*Fields of View, Image Quality*) and processing (*Fish Detection* and *Species Recognition Errors*).	Users	Latest Stage	Ground-Truth evaluation of *Fish Misdetection* over *Image Quality* (*Video Analysis* tab)
Data Interpretation				
Error-prone conditions in Computer Vision Results	Errors in computer vision results may be extrapolated from errors measured in test conditions, compared to the conditions specific to computer vision result subsets (*Image Quality, Field of View*).	Users, Experts	Mid-Term	Classification of videos and fish appearance, Filtering options (*Visualization* tabs)

Table 2. The uncertainty factors, identified by users all along design stages and complemented by system experts.

“Graham (aged 8)” Memory box demo

Graham Cooper
University of Lincoln, UK
3 Reed Lane, Stansted,
CM24 8GS, UK
jayumcooper@me.com
+447789 544552

ABSTRACT

“Graham (aged 8)” is an attempt to explore the notion of photographic memory, or lack there of. A demo installation, in the form of a Victorian viewing box, housing a digital display. The author presents this work as a personal journey, an autobiographical undertaking, which formed part of his research and practice during his Masters study in Digital Imaging and Photography. The work as a result could be interpreted as a viewpoint of one individual, but the intention is that through engaging with the piece viewers are encouraged to reflect on their own histories and archival habits.

Author Keywords

digital imaging; archiving; work in progress; conference demo;

ACM Classification Keywords

H.5.2. Prototyping; H.5.m. Miscellaneous;

“GRAHAM (AGED 8)”

My childhood was heavily documented through amateur photography, with album after album of prints displaying my journey from birth to university graduation. I can recall all of these images now with no need to return to the original source. I struggle however to recall any of these moments taking place. I have no actual memory of meeting Mickey Mouse, but I have a photographic record that tells me otherwise and this is true of all other occasions captured from the momentous to the mundane.

It's only when I tasked myself to recall distinct moments of my past that 8 moments, 8 scenes and locations arose. There are no photographs of these locations in the albums, and yet I can remember these places, these moments. Are these actual memories? Without evidence how can I be sure?

Presented as a viewing box (*Fig.1&2*) designed and constructed by my father and myself, individual viewers are granted permission to look inside where they are confronted with a series of 8 images that depict the only remaining memories of my childhood. The validity and truth of these scenes are presented as a question through digital manipulation. The work is accompanied by a collection of annotated sketches (*Fig.3*) that talk of change, loss, and reflections on the past. The box is an empty space, absent, trying to map / diagram this phenomenon of capturing memories, prompting us to reflect on the basic paradox of archival culture, archive fever, where we become appendages to the archive, our purpose in life is to produce, record and store.

I’d like to use this opportunity to demonstrate this piece as an intervention into my practice, affecting how I develop the piece further. The piece was created through catharsis, and as a result is incredibly personal, and by sharing this work with an audience the main goal is to gain feedback on the work itself, but to also to trigger thought and conversation on the subject of archival memory procedures.

Figure 1. The Memory Box

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

With the current climate of digital capture of the fantastic to the mundane and everyday, we are all creating our own memory archives, and with the prevalence of 'sharing' we are adding to the global record of human existence, but through this act are we losing our attachment to our own narratives? In my case, the analogue process of the family album, this archival act, was decided for me, and has had an affect on my own ability to recall the events of my life. Are we now voluntarily sacrificing our memories, without real thought or consideration to the future?

Figure 2. The Memory Box Version 1.0

Figure 3. Drawn from memory, one of the sketches that accompany the installation.

The AffectButton: a Digital Self-report Tool for Emotion

Joost Broekens
Interactive Intelligence Group, TU Delft
Mekelweg 4, Delft, The Netherlands
Joost.broekens@gmail.com

ABSTRACT

The AffectButton is a self-report tool that enables users to report upon their emotion, mood or attitude about things. It is a simple interactive button that presents a dynamically changing facial expression. The expression generated by the AffectButton changes based on mouse movement insight the button and the user selects the expression that best matches his or her feeling. In doing so, the user in fact enters three values on the well-known affective dimensions Pleasure (Valence), Arousal and Dominance. The resulting values and the usability of the AffectButton have been extensively evaluated. Here we give a short summary of the motivation, findings and recent developments.

Author Keywords

Emotion, Affect, Self-report, Affective Labeling.

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION

Being able to gather emotional feedback from users is important for many reasons. First, some applications need emotional user feedback in order to function. For example, recommender systems that are based on product feedback in fact gather an affective appraisal of that product. Second, user studies that measure emotional impact of novel technology or products need valid and reliable means of doing so [1]. For example, what emotional experience is associated with a new car model [1]. Third, valid and reliable emotion feedback is important for psychological reasons, e.g., when systems are developed to keep track of elderly well-being. In general, valid, usable, and reliable measurement of emotion is key to better understanding the experience of people, and usage scenarios include: affective labeling of videos, music, and products, gathering feedback on performances and events, polling data on persons (e.g. popularity polls), and so on.

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

THE AFFECTBUTTON

The AffectButton (see Figure 1, 2) enables such measurement of affect. The AffectButton is a measurement instrument that enables a user to give detailed and subtle emotion feedback about his/her feelings, mood, and attitudes towards people and products. The user selects the best-matching facial expression that goes together with his or her feeling, and then clicks. After clicking the AffectButton generates 3 values, one for each of the affective dimensions Pleasure (positive versus negativeness of the feeling), Arousal (active versus passive feeling) and Dominance (dominant versus submissive feeling) [2], in a similar way as does the Self-Assessment-Manikin [3].

Figure 1: AffectButton and its prototypical extreme states.

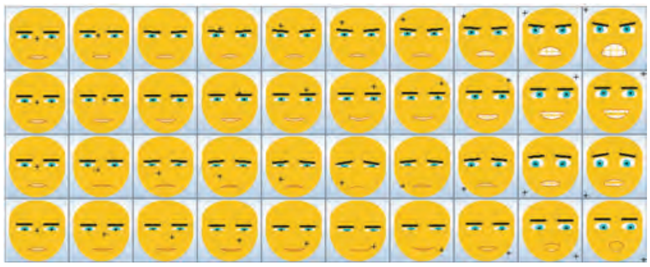

Figure 2: AffectButton faces corresponding to different locations of the mouse (+marker) inside the AffectButton.

EVALUATION AND USAGE

The AffectButton has been evaluated in a wide variety of usage settings including real time affective labeling of music [4], self-report of user mood, labeling of emotional stories and words [5] and affective labeling of holidays [6]. Research also shows that when users have a strong opinion, they like the additional detail that can be given with multiple affective dimensions in addition to thumbs up/down or 5-point star rating system [6]. Further, a large variety of test subjects has by now successfully used the AffectButton [5].

NOVEL DEVELOPMENTS: VIDEO LABELLING

A recent advancement in facilitating and testing the use of the AffectButton is the creation of an online tool for the labelling of affect. The tool shows how straightforward rating of video content on three affective dimensions becomes when using the HTML5 implementation of the AffectButton. The button is coupled to the video time, and when it is clicked, an affective timestamp is generated. This allows, for example, affective annotation of video (and audio) over time, generating affective traces of the media. For a link see www.joostbroekens.com.

SUMMARY

Compared to other methods of self-report, the AffectButton enables one-click three dimensional affective input, is easy to use with relatively little explanation, provides valid and reliable output in a format that is easy to aggregate making it a useful instrument to gather affective information from large groups of people. For a review of existing methods, as well as an in depth discussion of the AffectButton workings, its potential and limitations, please see [5]. For an interactive (demo) version please see www.joostbroekens.com. Also see the previous link for downloadable open source implementations of the AffectButton available in HTML5, Python and Java.

REFERENCES

1. Desmet, P.M.A. and P. Hekkert, “Framework of product experience.” *International Journal of Design*, (2007). 1(1): p. 57-66.
2. Mehrabian, A., *Basic Dimensions for a General Psychological Theory*, (1980) OG&H Publishers.
3. Bradley, M.M. and P.J. Lang, “Measuring emotion: the Self-Assessment Manikin and the Semantic Differential.” *Journal of Behaviour Therapy and Experimental Psychiatry*, (1994) 25: p. 49-59.
4. Broekens, J., A. Pronker, and M. Neuteboom, “Real time labeling of affect in music using the affectbutton.” In *ACM Workshop on Affective Interaction in Natural Environments*, (2010) ACM. p. 21-26.
5. Broekens, J. and W.-P. Brinkman, “AffectButton: A method for reliable and valid affective self-report.” *International Journal of Human-Computer Studies*, (2013) 71(6): p. 641-667.
6. Broekens, J., et al., “Factors Influencing User Motivation for Giving Online Preference Feedback.” In *5th Multidisciplinary Workshop on Advances in Preference Handling (MPREF'10)*, (2010).

ReSearchNow – Playful App for Participation and Engagement in (Design) Education

Marije ten Brink
Brinka / HvA
Amsterdam
marije@brinka.nl
m.ten.brink@hva.nl

ABSTRACT

The mobile app ReSearchNow stimulates alert observation and abstract thinking. The pilot version is used in (design) education, offering students a richer learning experience. Students are triggered to participate and add content to the app, inside and outside the classroom and during and outside scheduled study timetables. Students engage in topics related to the study program via the app and are motivated to add content by playful interaction and principles of game play.

The usage of and feedback on the pilot version of ReSearchNow by a group of ± 180 students makes clear that ReSearchNow has potential to grow. The majority of the students agreed with the statements: 'The app helped me to engage in the study assignment' and 'Using the app was a enjoyable way to engage in my study'.

Author Keywords

Social; interactive tools; education; participation; playful interaction; game play; serious games; sharing; co-creation; association; meaning

ACM Classification Keywords

K.3.1 [Computers and Education]: Computer Uses in Education---collaborative learning; K.8.0 [Personal Computing]: General---Games; H.5.0 [Information Interfaces and Presentations (HCI)]: General; H.3.1 [Information Storage and Retrieval]: Content Analysis and Indexing; H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval; H.2.8 [Database Management]: Database Applications---image databases.

General terms

Design, Experimentation

INTRODUCTION

Interactive (mobile) tools are used more and more in an educational context. Often they serve as a means to support or improve communication between student and tutor, for example by using blogs to deliver study results or a Facebook group to share updates and interesting facts. These tools offer support, but are not aiming on deepening knowledge of a specific curriculum.

Also, digital educational content is made to suit a specific student population, for example a module to learn a foreign language. In a step-by-step approach the student is guided through an educational module. These kind of applications offer deepening of specific knowledge, but miss out on interaction and dynamics with the content itself and are not employable for other or broader contexts. The same is true of MOOC's (Massive Open Online Courses).

I am interested in developing and applying interactive tools in an educational setting to improve participation and involvement of students with the content of curricula, by using playful interaction and principles of game play.

EDUCATIONAL VALUE

The educational value of ReSearchNow lies in the ability of students to associate an image to a textual concept. This transformation from text to image requires abstract thinking and creativity; within a restricted time frame the student has to form a mental model of the concept and scan his environment for visual clues that link to that model. In (design) education 'learn to observe' and abstract thinking are crucial skills for all students. This includes the ability to assess (design) work of fellow students and understand the relation of the work to the intentions of the maker. Being immersed in translation and association of text to image and vice versa is at the core of ReSearchNow.

The app ReSearchNow involves students in- and outside the classroom. Since the notifications arrive on different moments every day, students will not always be in the classroom or even at the University. This 'flipped classroom' principle is under study a lot the past few years [1], because of its potential to involve students more and deeper with educational content and focus.

ReSearchNow is used individually, but results in co-creation with fellow students. Seeing how others visually associate a textual concept can be inspiring and infor-mative. Using the app in a smaller community of people recognizing each other, such as students in a (design) educational setting, adds personal meaning and makes peer-to-peer assessment possible.

ReSearchNow (the tool) itself is also object of study. While experiencing the tool on a daily basis students are invited to

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

think about ways to make the tool more relevant, playful and/or usable, or to come up with ways to distribute the content of ReSearchNow via existing or novel media channels in new contexts.

SOCIAL CONSTRUCTIVISM

Constructivism suggests that learning is the process of adjusting our own understanding of the world around us through reflection on our experiences [2]. The world cannot be known directly, but rather by the construction imposed on it by the mind. The concept of constructivism is relevant to ReSearchNow, but within the overall con-structivist family, there is the position of the social constructivists that is especially interesting in relation to ReSearchNow. Social constructivists recognize that influen-ces on individual construction are derived from and preceded by social relationships. Constructivism posits a highly individualistic approach without reference to social interaction, contexts, and discourses, while social con-structivists move to more social explanations [3].

ADDED VALUE OF GAME PLAY

Grasser, Chipman, Leeming and Biedenbach [4] list in their paper 'Deep learning, emotions and games' a total of eight suggestions that indicate why games are succesful psychologically. Five of them apply to ReSearchNow:

1. Arousal of interest, Challenge and Fantasy

Arousal of interest in a specific concept is the very aim of ReSearchNow, by offering students a tool to visually research the meaning of that concept. This feature is *endogenous* to the app, rather than just a frivolous aspect of it that is *exogenous* to the concept.

2. Play

ResearchNow has the potential of integrating study and play. The app appeals to identity, imagination and self, while at the same time doesn't look similar to formal education.

3. Challenge and the Goldilocks Principle

Good games are not too hard or too easy, but just right (i.e. the Goldilocks principle) and at the zone of proximal development or at the brink of other zones of ability, cognition and emotion. In ReSearchNow the challenge is mainly set by peers through their ability to 'wow' a picture or not. Different levels of complexity and ambiguity of the daily notifications (the concepts) add another challenge to the experience.

4. Feedback

Peer-feedback on performance is the main motivational drive in ReSearchNow at this moment.

5. Types of interest

ReSearchNow has the potential to cater both situational interest – because of the immediate response being asked by the app, every day and within a certain time frame – as well as individual interest because of the personal ambition to create a recognizable and meaningful series of pictures.

HOW DOES RESEARCHNOW WORK?

ReSearchNow sends out a daily push notification with a meaningful word or concept. On a different moment every day somewhere between 8:00 AM and 22:00 PM students are prompted to make a picture of the topic-of-the-day and add it to the app. The time period to make a picture is restricted to one hour. After that time period it is not possible anymore to make a picture and the student has to wait for the next opportunity the next day. Pictures are added to the app immediately and they are visible to all other members of the student group. The pictures of all students together are grouped in a mosaic and form a collective photographic association, a shared visual view of a concept.

Students can 'wow' ('like') pictures of other students, making it possible for individuals to stand out from the rest and be recognized for it.

When tapping a picture in the mosaic the picture is shown in detail with automatic generated metadata: topic, name photographer, location, time and date. This metadata is tappable and leads the user to a new mosaic, revealing other visual patterns and meaning.

Future ideas evolve around: being involved in setting the topic-of-the-day, 'follow' fellow students with interesting content, adding tags or other content to increase the meaningfulness of the photo in relation to the concept, reaching other levels in the app by being the first to add a picture to a topic or by collecting the most 'wow's, battles between individual students or between groups. These are all examples of playful interaction and/or game mechanics, aiming on increasing participation, involvement and what Csikszentmihaly calls the *flow* experience [5].

STATUS, PILOT VERSION AND USER GROUP

The pilot version of ReSearchNow is launched on January 3rd 2014. A group of approximately 180 students of the study Communication and Multimedia Design (CMD) at the Amsterdam University of Applied Science (HvA) in Amsterdam, the Netherlands, started using the app in the 'Design for Interaction' mandatory course in the 3rd year of study. Additionally, a small group of tutors and other people interested were added to the user group.

ReSearchNow is developed as a 'webapp' in order to give all students, with all types of phones and tablets, the opportunity to work with the app, while at the same time keeping the time and money spent to an acceptable and realistic level. The technical implementation of the app is handled by three students in the program and an external front end developer. Notifications are dealt with via a third party app: Jeapie. The daily concepts can be defined and scheduled up front via a admin-environment.

The pilot version of ReSearchNow has basic functionalities and design and the implementation leaves room for improvement. For example: the log in procedure and the notifications via Jeapie were not optimal. Also the ability to upload your own profile picture and to have clear insight in the amount of 'wow's' given and received were not implemented yet. Even so, the majority of the students agreed with the following statements: 'The app helped me to engage in the study assignment' (65%) and 'Using the app was an enjoyable way to engage in my study' (70%).

After the trial with the pilot version of ResearchNow a questionnaire is sent to all participating students. The questionnaire consists of four parts: Name, group and used phone or tablet (1); Handling the daily notifications (2); Adding content to the app (3); Agree/disagree with statements about the usage of the app. The outcome still has to be examined in detail.

CONCLUSION

After working with a pilot version of the app in January 2014 with a group of ± 180 students the question whether the use of ReSearchNow has motivated participation and engagement in the curriculum and therefore facilitates a richer learning experience can be answered with a modest 'yes'. Although the pilot version of ReSearchNow leaves room for improvement, the majority of the students agreed with the following statements: 'The app helped me to engage in the study assignment' (65%) and 'Using the app was an enjoyable way to engage in my study' (70%). This gives reason, to believe that ReSearchNow has potential to grow. More research, prototyping and testing has to be done.

REFERENCES

1. Gray Wilson, Stephanie. A Method to Address the Challenges of an Undergraduate Statistics Course. In *Teaching of Psychology* 40, London: Sage (2013).
2. Akers, R. "Web Discussion Forums in Teaching and Learning", *The Technology Source Archives*, 1997. http://technologysource.org/article/web_discussion_forums_in_teaching_and_learning/ (accessed March 16, 2014).
3. Young, Richard A., Collin, A. Introduction: Constructivism and social constructionism in the career field. In *Journal of Vocational Behavior* 64, Elsevier (2004), 373-388.
4. Graesser, A., Chipman, P., Leeming, F. and Biedenbach, S. Deep learning and emotion in serious games. In *Serious Games: mechanisms and effects*, Routledge (2009), 83-102.
5. Csikszentmihalyi, Mihaly. *Beyond boredom and anxiety: experiencing flow in work and play.* San Francisco: Jossey-Bass Publishers (2000).

Chi sparks 2014

Gamesourcing Expert Painting Annotations

Myriam C. Traub
Centrum Wiskunde & Informatica
Science Park 123, Amsterdam
myriam.traub@cwi.nl

Jacco van Ossenbruggen
Centrum Wiskunde & Informatica
Science Park 123, Amsterdam
jacco.van.Ossenbruggen@cwi.nl

Jiyin He
Centrum Wiskunde & Informatica
Science Park 123, Amsterdam
jiyinhe@gmail.com

Lynda Hardman
Centrum Wiskunde & Informatica
Science Park 123, Amsterdam
lynda.hardman@cwi.nl

ABSTRACT

Online collections provided by museums are increasingly opened for contributions from users outside the museum. These initiatives are mostly targeted at obtaining tags describing aspects of artworks that are common knowledge. This does not require the contributors to have specific skills or knowledge. Museums, however, are also interested in obtaining very specific information about the subject matter of their artworks. We present a game that can help to collect expert knowledge by enabling non-expert users to perform an expert annotation task. This is achieved by simplifying the expert task and providing a sufficient level of annotation support to the users. In a user study we could prove the usefulness of our approach.

Author Keywords

Crowdsourcing; expert task; annotation; wisdom of the crowd.

ACM Classification Keywords

H.5.2. User Interfaces; H.5.3. Group and Organization Interfaces; K.4.3. Organizational Impacts

INTRODUCTION

The Rijksmuseum Amsterdam makes large parts of its collection available online and continuously adds new items. Visitors of the website cannot only look at the images, they are also invited to interact with the items by annotating or downloading them in high resolution to use them in a creative way [4]. The annotations that users contributed so far, however, show that mechanisms are needed to detect and remove incorrect annotations and to help users provide more qualitative annotations.

The SEALINCMedia project [5] aims to support the Rijksmuseum Amsterdam in this endeavor. We develop an online platform called Accurator that combines functionalities to manage the administration and the execution of the annotation process. The predominantly art-historical knowledge available from museum professionals is extended by involving people from outside the museum.

To maximize not only the coverage, but also the level of detail of annotations, we collect common knowledge as well as expert knowledge. Common knowledge, i.e. information that can be provided by any user (e.g., “the painting shows a blue bird and a branch with red leaves”) can be collected via crowdsourcing platforms without knowing the skills of the users. Obtaining expert knowledge, (e.g., “the bird depicted is a Daurian Redstart”) from experts is in comparison much more difficult. Finding experts of the targeted niche (i.e. ornithologists), persuading them to help in an annotation task and keeping them engaged are considerable challenges. While overcoming these difficulties promises to lead to high quality annotations we decided not to solely rely on the traceability and good will of experts.

Therefore, to bridge the gap between the availability of lay people on crowdsourcing platforms and the scarcity of expert knowledge, we have developed a game that transforms an expert annotation task in a way that it can be carried out by non-experts. This adds another source for expert knowledge to our platform (see Fig. 1) and thereby enhances our chances of obtaining qualitative and specific descriptions of paintings.

We conducted a study [2] to investigate whether our approach is suitable to support crowdsourcing of expert annotations. We investigated

- whether the task is actually feasible for non-experts at a level that is comparable to experts;
- whether non-experts improve while playing;
- how the partial absence of the correct answer influences the users’ performance; and
- whether the aggregation of user judgements improves the agreement with experts.

During the conference, we will give conference participants the chance to test their knowledge of art-history with the “imperfect” version of the game and enjoy the competition with others while – hopefully – learning more about classification of paintings.

THE ART GAME

Our game is based on the online tagging game used for the Fish4Knowledge project [1]. For comparability reasons it was largely left unchanged. The game is available online [6].

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

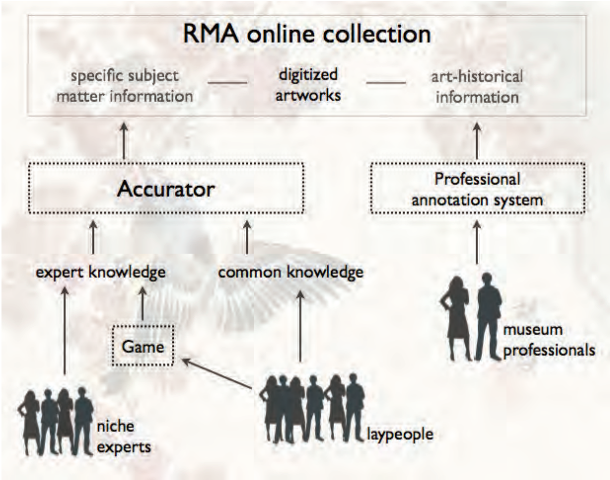

Figure 1: How laypeople and experts can contribute to a more detailed and varied description of artworks.

The Task

The task we selected for the game is the classification of paintings into 17 selected subject types (see Table 1) from the Art & Architecture Thesaurus (AAT) [7]. This task is usually performed by museum professionals. We chose the subject type classification because the collection items of the Rijksmuseum Amsterdam lack this information and because we were able to re-use data collected for a graduation project [3].

The Interface

On the start page, users are given instructions on how to perform the annotation task (if needed, they can go back to the instructions at any point in the game). Users are asked to be as *specific* as possible in their judgements. This means, that when they can choose between a general subject type (e.g. *figures*), and a specific subject type (e.g. *half figures*), they should select the more specific one.

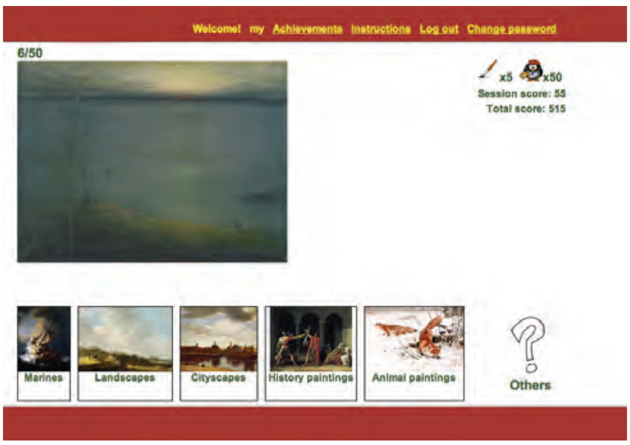

Figure 2: Interface of the art game with the large query image on the upper left. The five candidate subject types are shown below, together with the *others* candidate.

After creating an account, users can start playing the game. They are presented a succession of images (referred to as *query images*) of paintings (see Figure 2) taken from the Steve Tagger [8] data set. Apart from the image, we provide no further information about the painting. Within the first ten images that are presented to the user, there are no repetitions. Afterwards, images may be repeated with a 50% chance.

The query images are presented prominently in the upper left corner of the interface. Below, we present a pre-selection of six *candidates*. Five of these candidates represent subject types and one of them (labeled “others”) can be used if the assumed correct subject type is not presented. A candidate consists of an image, a label (AAT subject type) and a description. For each subject type we selected one representative image from the corresponding Wikipedia page, e.g., [9].

These images are intended to give users a first visual indication of which subject type might qualify and it makes it easier for users to remember it. If further information is needed to judge the image, users can display short descriptions taken from the AAT by moving the mouse cursor over the candidate, for example:

Marines*:*

“Creative works that depict scenes having to do with ships, shipbuilding, or harbors. For creative works depicting the ocean or other large body of water where the water itself dominates the scene, use ‘seascapes’.” [10]

The descriptions of the subject types are important, as the differences between some subject types are subtle.

Subject type	#
full-length figures	40
landscapes	33
half figures	13
allegories, history paintings, portraits, animal paintings, genre, kacho, figures	8
townscapes	6
flower pieces	5
marines, cityscapes, maesta, seascapes, still lifes	3

Table 1: Used subject types and the number of expert annotations.

The Feedback for Users

To motivate users to annotate images correctly and to give them feedback about the "correctness"[1] of their judgements, they are awarded ten points for correct choices and one point for the attempt (even if incorrect). On the top right corner of the interface, users can see their current and overall score.

After finishing a round of 50 images, users are directed to the dashboard, where they can compare their scores to those of the other players.

The feedback for the users is based on a comparison of their judgements to annotations given by experts of the Rijksmuseum Amsterdam collected by [3]. From this data set, we selected 168 expert annotations for 125 paintings (Table 1). The number of annotations per painting ranges from four (for one painting) to one (for 83 paintings). These multiple classifications are considered correct: a painting showing an everyday scene on a beach [11] can be classified as *seascapes*, *genre*, *full-length figure* and *landscapes*.

EVALUATION

We evaluated the suitability of our approach with two experiments. The first experiment simulated "perfect" data, meaning that the correct candidate was always presented to the user. The results of this showed us that the users are able to perform the task and improve over time. We used this data as a baseline to compare the results of the second experiment to. Here, we deliberately removed the correct candidate in 25% of the cases to simulate an imperfect data set. The analysis of the results showed that while the agreement between users and experts is higher in the first experiment, it is still acceptably high in the second setting. The learning performance, however, was much lower, which suggests that users should be given a training phase before they can successfully play on imperfect data.

The application of majority vote on the users' judgements leads to a noticeable improvement of the agreement with the experts. For some paintings, however, the agreement remained very low. We consulted another expert from the Rijksmuseum Amsterdam and identified two main reasons: The classification done by the experts was either incomplete or incorrect or the correct classification was only possible if context information about the painting was known.

The setup of the experiments, the analysis and our findings are described in detail in [2].

CONCLUSIONS

We present a game for classifying paintings into categories of a professional vocabulary. We have simplified the task by reducing the number of categories and added assistance for the users to perform the task. In a user study we could show that non-expert users are able to perform the task and improve over time. A "Wisdom of the Crowd" effect was observed when we aggregated the users' votes. In some cases of very low agreement with experts we could identify incompleteness or incorrectness on the expert's side to be a reason.

This approach allows us to enlarge the circle of potential contributors of high quality annotations.

REFERENCES

1. Jiyin He, Jacco van Ossenbruggen, and Arjen P. de Vries. "Do you need experts in the crowd? : a case study in image annotation for marine biology." In *Proceedings of the 10th Conference on Open Research Areas in Information Retrieval*, OAIR '13, pages 57–60, Paris, France, France, (2013). Le Centre De Hautes Etudes Internationales D'Informatique Documentaire.
2. Myriam C. Traub, Jacco van Ossenbruggen, Jiyin He, and Lynda Hardman. "Measuring the Effectiveness of Gamesourcing Expert Oil Painting Annotations." In *Proceedings of the 36th European Conference on Information Retrieval*, volume 8416 of Advances in Information Retrieval, pages 112–123. Springer International Publishing Switzerland, (2014.)
3. Sjaak Wouters. Semi-automatic annotation of artworks using crowdsourcing. Master's thesis, Vrije Universiteit Amsterdam, The Netherlands, (2012).
4. https://www.rijksmuseum.nl/en/rijksstudio
5. http://sealincmedia.wordpress.com
6. http://sealincmedia.project.cwi.nl/artgame/accounts/login/
7. http://www.getty.edu/research/tools/vocabularies/aat/
8. http://tagger.steve.museum/
9. http://en.wikipedia.org/wiki/Maesta
10. http://www.getty.edu/vow/AATFullDisplay?find=marines&logic=AND&note=english%3DN&subjectid=300235692
11. http://tagger.steve.museum/steve/object/280

[1] By "correct" we mean that a given judgement is in line with the expert.

chi sparks 2014

The Programming Language as Human Interface

Steven Pemberton
CWI
Science Park 123, NL-1098 XG Amsterdam
steven.pemberton@cwi.nl
http://www.cwi.nl/~steven

ABSTRACT

Programming languages are mostly not designed for humans, but for computers. As a result, programming time is increased by the necessity for programmers to translate problem description into a step-wise method of solving the problem. This demonstration shows a step towards producing more human-oriented programming languages, by developing an interactive map application in a language that allows specification of *what* needs to be solved rather than *how* to solve it.

Author Keywords

Programming languages; HCI.

ACM Classification Keywords

F.3.m Studies of Program Constructs; Miscellaneous; H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION

Despite all appearances to the contrary (active at night, little-to-no sleep, diet consisting solely of pizza, sugar and caffeine) programmers are humans too. This might lead you to conclude that programming languages must be a human-computer interface, and that therefore they would be designed as a human interface, with requirements analysis, iterative design and user testing. But with notable exceptions, e.g. [1], most programming languages are not designed like this.

HISTORY

In the 1950's, when computing seriously started, computers were expensive, so expensive in fact (in the millions) that you seldom bought a computer but leased one instead. Often you would get a few programmers for free as a sweetener for the deal; in other words, compared to the price of the computer, programmers were essentially free. Partly for that reason, it was essential that programmers made their programs as efficient as possible: within bounds it didn't matter how much time you spent on programming, as long as you reduced the load on the computer. The first programming languages were designed around this time, and consequently they all had the basic premise that you had to tell the computer what to do: the programmer had to turn the problem specification into a step-by-step solution for that problem, with the aim of reducing the time the computer had to spend on it, and not reducing the time the programmer had to spend on it.

Now, 60 years later and the tables are turned: computers are more or less free compared with the cost of the programmer. But whereas computers are now millions of times faster than the computers of the 50's, programmers are still programming with languages that are visible descendants of the original programming languages. Programmers are still telling computers what to do, and as a result are barely more productive than 60 years ago.

HCI AND PROGRAMMING

ISO 9241 [2] defines usability as the effectiveness, efficiency and satisfaction with which users achieve their goals in a particular environment.

In a major study of the costs of programming that eventually led to the design of the Ada programming language, the United States Department of Defense discovered that 90% of the cost of producing software was in debugging. In another piece of research by IBM [3], over a large range of software, both in terms of size, and in terms of programming languages used, it was discovered that the number of bugs in a program did not grow linearly with the length of the program, but as a power function

$$b \propto s^{1.5}$$

In other words, a program 10 times longer has around 30 times more bugs, or alternatively, a program of one-tenth the size costs 3% of the larger program.

What this suggests for language design is that the efficiency part of the HCI equation would best be met by programming languages that are designed to be as compact as possible.

ABC

In earlier work, the author was involved with the design of a programming language for beginners [1]. Although this work was done before the term HCI was in common use, it was designed with what would now be recognised as classical HCI techniques: requirements analysis, iterative design, and user testing. The resulting language ABC turned out to be more powerful than was originally foreseen, being

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

useful for 'real' programming as well, and not just for beginners (and in fact went on to become the basis for the programming language Python [4]).

Although the language used mostly classical programming and control structures, such as assignment, procedure and function calls, if and while statements and so on, experience showed that programmers were around an order of magnitude faster at programming than with the classical programming languages it was compared with, such as Pascal, Basic, or C [5]. The main reason behind this was the use in ABC of a small number of high-level data structures. Most programming languages provide a set of low-level data structures, which may then be used to design and build higher-level structures. A key realisation with the third iteration of ABC was that it was the high-level data structures that the programmers needed, and they hardly ever used the low-level structures except to build the higher-level ones. An analysis of programming structures in programming led to a set of 5 data types in ABC that provided as a primitive the essential data structuring that programmers really needed.

XFORMS

XForms [6] is a programming language that investigates another aspect of programming: the control structures. XForms was originally designed (as the name suggests) for Forms applications, but in its second iteration became generalised so that more general applications could be written with it as well.

XForms is unusual in that many of the administrative tasks normally associated with programming are left to the computer to solve. The program is stated far more in terms of *what* needs to be solved than *how* to solve it. As a result, program size, and therefore programming time, is sharply reduced. Experience with several projects shows an order of magnitude reduction in project times.

This demonstration develops an interactive map application in XForms [7], ending up with a functional, working application in around 150 lines of code, an application that would normally require tens of thousands of lines in a traditional language such as Javascript. An interesting property of the resultant program is that it doesn't contain a single while loop (in fact it couldn't, because such a thing doesn't exist in XForms).

A NOTE ON NOTATION

XForms is expressed in XML. This is principally to allow it to be integrated with diverse other XML languages such as XHTML and SVG. However, this is purely a notational issue. Whether you write

```
integer i;
```

or

```
i: integer;
```

or

```
<bind ref=”i” type=”integer”/>
```

makes no difference to the concepts being described.

A TASTE OF THE PROGRAM

XForms does not normally work in a “first do this, then do that” style of programming that most languages use, but specifies relationships between data that the system autonomously keeps up to date.

For instance, if you have several pieces of data

```
site: http://tiles.osm.org/
zoom: 10
x: 526
y: 336
```

and specify a relationship:

```
url= site+zoom+”/”+x+“/”+y+”.png”
```

then the URL will always be kept up to date if the underlying data changes. (Note that this is not an assignment in traditional programming terms, but a unidirectional invariant relation).

Furthermore, if you choose to output the image connected to the URL:

```
output(url, “image/*”)
```

then the image on the screen will change as the data does.

If you then take a position on the world's surface in the coordinate system used by the map server:

```
posx: 34477602
posy: 22058667
```

you can then specify the relationship between this position and the map tile that that location is on:

```
x=floor(posx ÷ scale)
y=floor(posy ÷ scale)
scale=2^(26-zoom)
```

(since these are invariants, the ordering doesn't matter; 26 in this case is a function of the tile size and the maximum value of zoom, so that scale represents the number of locations on a tile at any particular level of zoom).

Hence, any time the position gets updated, so does the URL, and so does the display of the corresponding map tile.

Similarly, if the value of site gets changed to another tile server (for instance to a server that serves maps in a different style) as long as the server uses the same coordinate system then it is a trivial issue to display a map in that different style.

CONCLUSION

XForms programming requires a different style of programming than with traditional languages, which can be sensibly compared with how spreadsheets work.

Experience has shown that such a style of programming greatly reduces programming time. One very-large-scale project that tested XForms in a process that a company had done many times before reduced the time and staffing from five years with a team of thirty to one year with a team of ten, in other words from 150 person years to 30; another company that translated a large collection of Javascript programs to XForms reported that the resultant XForm programs were around a quarter of the length of the corresponding Javascript (which plugging into the equation quoted above would lead you to expect a saving of programming time of around a factor of 8).

The program presented here is itself a mere 150 or so lines of code, which is an order or two magnitude less than the equivalent code that would be needed in a traditional language like Javascript.

REFERENCES

1. Geurts, Meertens, Pemberton, The ABC Programmer's Handbook, Prentice-Hall, 1985.
2. ISO 9241-11:1998 Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability. http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=16883
3. Brooks, FP, The Mythical Man-Month. Addison-Wesley. ISBN 0-201-00650-2, 1975.
4. Bruce Stewart, An Interview with Guido van Rossum, ONLamp.com, http://www.onlamp.com/lpt/a/2431, 2002.
5. Steven Pemberton, An Alternative Simple Language and Environment for PCs, IEEE Software, Vol 4, No. 1, pp 14-22, Jan 1987.
6. John M. Boyer (ed.), XForms 1.1, W3C 2009, http://www.w3.org/TR/xforms/
7. Steven Pemberton, A Map Application with XForms, http://www.youtube.com/watch?v=2yYY7GJAbOo, 2014.

Chi sparks 2014

REHAP – Modular and Interactive Rehabilitation Tiles

Victor Donker[1] **Bert Bongers**[2] **Panos Markopoulos**[1]

[1] Department of Industrial Design
University of Technology Eindhoven
Den Dolech 2, 5612 AZ Eindhoven, The Netherlands
victor@victordonker.nl, p.markopoulos@tue.nl

[2] Interactivation Studio,
Faculty of Design,
Architecture and Building
University of Technology, Sydney
bertbon@xs4all.nl

ABSTRACT

This paper describes the vision behind and development of a modular and interactive tile system, designed for patients suffering from a stroke. Traditional rehabilitation techniques are mainly based on mechanical structures and passive materials and are used in an improvising way by physiotherapists. On the other hand we see large-scale products equipped with a big amount of sensors and possibilities, however these are mostly complicated for both therapists and patients. In collaboration with medical researchers and practitioners a vision for individual and personalisable products has been formed. Observing and understanding the current situation in a stroke unit in a hospital in Sydney showed the strengths and on the other hand the needs of practitioners. Responding to this a modular and interactive approach was opted, which allows therapists to create exercises for single patients. Essential factors for physical rehabilitation are *motivation, customization and independence*. By using techniques like force sensitive resistors, magnetic and spring-loaded connections [1] and 3D printing, several sets of pressure sensitive floor tiles and a corresponding graphical interface with visual feedback have been produced, which are demo versions and frequently used as in-situ prototypes in hospitals in rehabilitation practices.

Author Keywords

Stroke rehabilitation, modularity, motivation, interactive, individualization

ACM Classification Keywords

H.5.2 User Interfaces: User-centered design

INTRODUCTION

In the world of medical and physical rehabilitation exercise techniques we can define two product types, the analogue and the products fully equipped with sensors. In the case of medical rehabilitation of stroke patients it is essential to keep track of the exercise results of the recovery process. The lack of ease to record data during physiotherapy exercises creates an opportunity for intelligent interaction design. During the last few years a vision and a variety of prototypes have been created with different design students and professionals [3,4,6]. The development and design of a modular pressure sensitive set of floor tiles, made for physical and medical stroke rehabilitation, will be described in this paper. The set of so-called REHAP tiles is displayed in Figure 1.

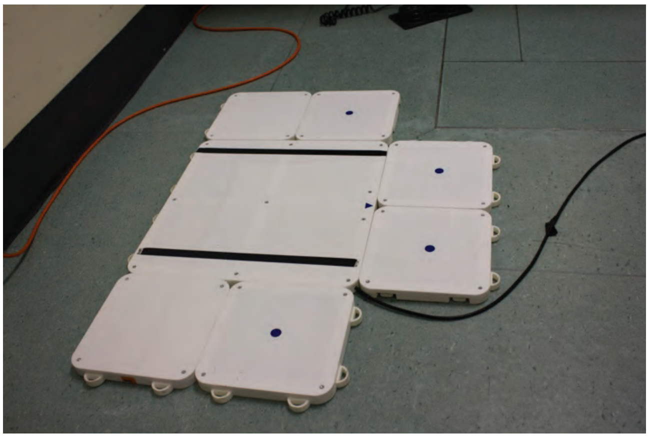

Figure 1. Tiles positioned in stroke unit in hospital.

APPROACH

The way of designing and developing has been according to a concept method developed by the first author, called OPTAR, which stands for Observe, Produce, Test, Analyze and Reproduce. By observing the situation it is possible to quickly see the practical environment, which does not only show problems but also strengths of the target group(s). This approach works inspiring, motivating and allows creating a vision and products that fit the needs of the target group(s) perfectly. Looking to problems and strengths, simple prototypes can be produced quickly that can solve issues for the chosen context. Prototypes will then be tested in the earlier observed environment. The tests and feedback will be analyzed, and from these results the prototypes can be reproduced. This cycle can be used iteratively, until a satisfying result has been produced. By working in this way, products can be developed that are thought out in detail. By frequently showing the end users possibilities (as prototypes, or 'provotypes' [2]) and involving them through out the process the initial ideas can turn into valuable product ideas, worth investigating and developing further.

Published in: van Leeuwen, JP, Stappers, PJ, Lamers, MH, Thissen, MJMR (Eds.) *Creating the Difference: Proceedings of the Chi Sparks 2014 Conference*, April 3, 2014, The Hague, The Netherlands.

MODULAR INTERACTIVE STEPPING TILES

In response to the need of the therapies, as explained in the approach section, we developed a modular set of balancing and stepping task tiles. The work was inspired by an exercise mat [9], which is a design adapted from the Dance Dance Revolution game paradigm. With this mat, with fixed sensors equipped in it, elderly or less mobile people can practice dancing and stepping exercises in the home environment. Digital games for physical therapy have been used successfully before [5].

The REHAP sensor tiles however are developed with the vision that it should perfectly fit the needs of an average stroke patient, and moreover, it should be personalisable for separate patients. Thanks to the modular approach a therapist can lay down an exercise floor for every single patient, based on the disability and goal of the training of that specific patient. The sensor tiles are fabricated using instant manufacturing, which creates variable possibilities to quickly develop possible prototypes to test in practice. The tiles are in total only 20mm high, which is far lower than other entertainment robotic tiles, among others like the Wii Balance Board. The REHAP tiles set consist of a main tile which is 400x400mm which is the center tile of the exercise, and measures the pressure of 4 points, the back and front of the left and right feet. Subsequently there is the possibility to connect small 'sub tiles' around the maintile. These tiles with one single force sensor are 200x200mm, and therefore there can be exactly 2 sub tiles connected at one side of the main, so in total around the maintile there could be 8 sub tiles. The purpose of the sub tiles is to allow therapists to set up stepping exercises for the patients, which is crucial for weight shifting, an important ability to be able to walk [8].

The design process was highly inspired by the practitioners, who suggested most of the key elements in the current design. Showing a continuous signal on a display of the pressure points has more value than only showing the balance right/left. For example the feedback is on a screen in front of the exercising patient, as looking forward is essential for a proper balance. This feedback is also of big advantage for the therapist, as it can be used as a communication and explanation tool, as shown in Figure 2.

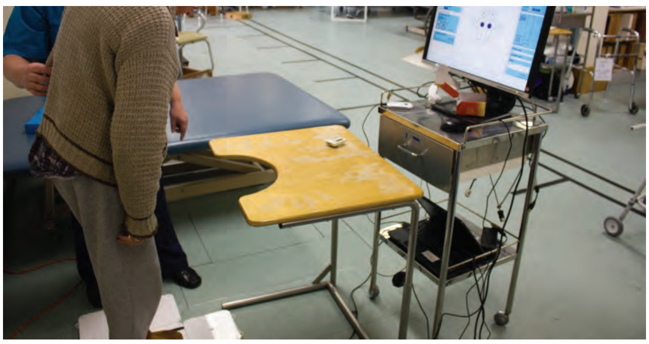

Figure 2. Patient exercising with therapist assistance.

The advantage of having a modular system [7] has proven when therapists asked if the tile could be used in a tilt table bed exercise, in which patients have to push themselves up in a bed by pressing against the footboard of the bed, as can be seen in Figure 3. Most patients were not aware of the behavior and strength of their legs and could not even believe that they actually had power in their leg to be able to press up. Getting feedback on the amount pressure on the back and front of both feet helps the patients to train the stability and strength in their legs. Another patient was not even with very motivating help from the therapist able to press up at all. He was not aware and didn't believe he had the ability and power to do so. When implementing the main tile and the GUI the patient saw a very little dot, which gave feedback of his own movement, allowing to believe in himself and to press even harder. The simple growing dot was of incredible motivating value for the patient.

Figure 3. Main tile being used in tilt table bed exercise.

GRAPHICAL USER INTERFACE

The Graphical User Interface (GUI) is build for the functions to control the settings of the exercise and overview, the live and direct feedback and a results section. A screenshot of the GUI, which is build in MAX/MSP, can be seen in Figure 4. As mentioned the interface is an information and communication system for both the patient and the therapist. The graphical illustration of growing dots based on increasing pressure has proven to be a simple yet very informative representation during balance and stepping exercises for stroke patients.

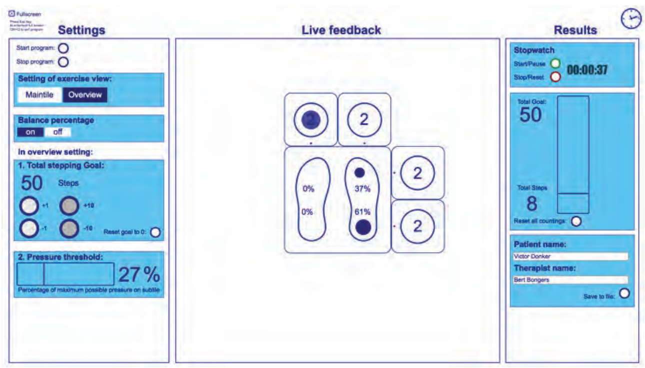

Figure 4. Graphical User Interface.

The interface corresponds directly to the layout of the physical tiles on the floor. When attaching a sub tile to the main tile the interface reacts to this and displays the attached tile directly on the screen. These sub tiles on screen show the amount of stepping repetitions on those single tiles, based on the pressure that has been put on them, when reaching the pressure threshold, which is to be controlled in the Settings part on the left of the screen. In the settings part the user can also choose to display percentages of the single pressure points, relatively to the complete pressure. Next to this the user can set the goal of amount of steps to be taken in total. The right part of the GUI shows a stopwatch and the progression of the total amount of steps toward the stepping goal. This is very valuable for therapists to have an overview of how the patient increases the average amount of steps per minute.

For logging purposes the right bottom part of the GUI also allows to fill in the patient and therapist name, with a possibility to write some notes of the exercise. When hitting the save button a .csv (comma separated value) file will be automatically produced, with information about that certain exercise. These .csv files are readable in Microsoft Excel and can be used to create graphical overviews of the process. Keeping an overview of the progression of a patient over a long time will help the guidance and will give a better view on if and how goals are achieved on long term.

CONCLUSION

Concluding from the development process and results we find that it is essential to go into the context early in the process to find the strengths and needs of the user. By observing the way of working, proposing ideas by bringing physical and quickly made ideas it is easier to test, analyze and discuss to find several key factors to design for. This frequent and iterative way of designing and including the users helped to create a modular balance and stepping exercise product that is currently used in the context where people benefit from the created values based on their needs. Values in this case include motivation, customization and independence. These values and the iterative OPTAR approach can be valuable for any design case and can be considered as valuable to pay attention to in the future for human-computer interaction. Especially creating individual products and systems by using a modular approach is an essential value in the project looking to the fact the product has been proven to be very valuable for patients and therapists in practical context.

ACKNOWLEDGMENTS

We would like to thank among others especially the Centre for Contemporary Design Practices (Sydney), the George Institute for Global Health (Sydney), therapists and patients in the Bankstown-Lidcombe Hospital (Bankstown, Sydney), the User Centered Engineering Research Group (Eindhoven) and Blixembosch (Eindhoven).

REFERENCES

1. Bdeir, A. and Ullrich, T. Electronics as Material: littleBits. Proceedings of the TEI ’11 (2011) 341-344.
2. Boer, L., Donovan, J., Provotypes for Participatory Innovation. DIS 2012, June 11-15 2012, Newcastle, UK.
3. Bongers, A. J., Smith S. T., Donker, V., Pickrell. M. and Hall, R. Interactive Infrastructures – Physical Rehabilitation Modules for Pervasive Healthcare Technology. In: A. Holzinger, M. Ziefle & C. Röcker (Eds.) Pervasive Health – State of the Art and Beyond. Springer, 2014.
4. Bongers, A. J., Smith S. T., Pickrell. M.,Hall, R. and Donker, V. Interactivated Physical Rehabilitation Modules. Proceeding of the International ACM Creativity and Cognition Conference (2013).
5. Geurts, L., Abeele, V. vanden, Husson, J., Windey, F. Overveldt, M. van, Annema, J. H. and Desmet, S., Digital Games for Physical Therapy: Fulfilling the Need for Calibration and Adaptation. Proceedings of the TEI’11 conference (2011) 117 – 124.
6. Pickrell, M. and Bongers, A. J. Re-Ability Sleeve – A Modular Wearable Rehabilitation Interface System. Workshop paper SmartHealth, OzCHI 2011.
7. Richardson, B., K. Leydon, K., Fernström, M. and Paradiso, J. (2004) Z-Tiles: Building Blocks for Modular, Pressure-Sensing Floorspaces. In *Proc* CHI 2004, ACM press.
8. Sherrington, C., Whitney, J.C., Lord, S.R., Herbert, R,D., Cumming. R.G., and Close. J.C., Effective Exercise for the Prevention of Falls: A Systematic Review and Meta-Analysis. The American Geriatrics Society, 56 (12), (2008).
9. Smith S. T., Sherrington C., Schoene D., Studenski S. and Lord S. R. A novel Dance Dance Revolution system for in-home training of stepping ability in older adults. British Journal of Sports Medicine, 45(5), (2011), 441-445.

AUTHOR INDEX

Bakker, René 36
Beauxis-Aussalet, Emma 152
Bekker, Tilde 91
Bennis, Arthur 45
Bongers, Bert 169
Broekens, Joost 158
Bruijnes, Merijn 86
Brussee, Rogier 60
Coelho, Joanna 78
Cooper, Graham 104, 156
Cremers, Anita 113
de Bruijn, Arjen 133
de Haan, Geert 126
de Moor, Aldo 108
de Valk, Linda 91
Dirkx, Bart 149
Donker, Victor 169
Egerer, Daniel 24
Eggen, Berry 113
Emmen, Dirrik 100
Foster, Derek 50
Frost, Jeana 74
Geurtsen, Annika 142
Gómez-Maureira, Marcello 11
Grimes, Jesse 130
Guy, Amy 139
Hardman, Lynda 152, 163
He, Jiyin 163
Heintze, Katrin 50
Hekman, Erik 60
Hermsen, Sander 74
Heydra, Chris 19
Hogenboom, Antoine 145
Jansen, Reinier 19
Jeurens, Jasper 36
Kelia, Mike 31
Khan, Vassilis-Javed 65
Lamers, Maarten 7, 122
Lampropoulos, Georgios 100
Lawson, Shaun 50
Lenior, Dick 45
Linehan, Conor 50
Loke, Ben 31
Markopoulos, Panos 169
Noldus, Lucas 31
Oei, Swie 133
op den Akker, Rieks 86
Pemberton, Steven 166
Peters, Rifca 86
Quanjer, Arnold Jan 122
Renes, Reint Jan 74
Rouw, Jos 147
Schoonheyt, Maureen 50
Schraffenberger, Hanna 11
Smulders, Bobbie 142
Spink, Andrew 31
Stappers, Pieter Jan 7
ten Brink, Marije 160
Teunisse, Carolien 11
Thayne, Martyn 104
Thissen, Maarten 7
Traub, Myriam 163
Turksma, Arthur 147
Ugur, Sirvan 133
van Beukering, Alice 91
van de Aalst, Ruud 149
van de Langkruis, Anneke 130
van den Elst, John 65
van Dijk, Jelle 24
van Eekelen, Wouter 65
van Egmond, René 19
van Leeuwen, Jos 7
van Ossenbruggen, Jacco 163
van Turnhout, Koen 36
Verbeek, Fons 11, 78
Waardenburg, Thijs 60
Zoon, Hanna 113